NEW INSIGHTS & UPDATES ON THE MOLECULAR EPIDEMIOLOGY AND ANTIMICROBIAL RESISTANCE OF MRSA IN HUMANS IN THE WHOLE-GENOME SEQUENCING ERA

EDITED BY: David Coleman, Anna Shore, Richard Goering and Stefan Monecke

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEW INSIGHTS & UPDATES ON THE MOLECULAR EPIDEMIOLOGY AND ANTIMICROBIAL RESISTANCE OF MRSA IN HUMANS IN THE WHOLE-GENOME SEQUENCING ERA

Topic Editors:
David Coleman, Trinity College Dublin, Ireland
Anna Shore, Trinity College Dublin, Ireland
Richard Goering, Creighton University, United States
Stefan Monecke, Carl Gustav Carus Technische Universität, Germany
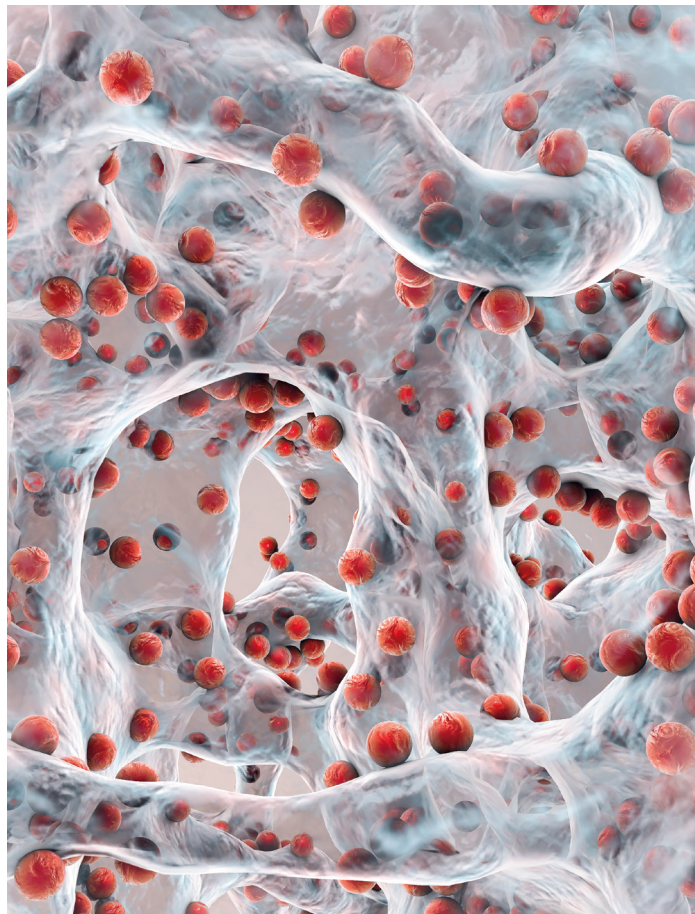
Image: Kateryna Kon/Shutterstock.com

Methicillin-resistant *Staphylococcus aureus* (MRSA) have been a major cause of healthcare-associated (HA) infection globally for several decades. During this time many distinct clones have emerged independently around the world, some of which have achieved pandemic status. More recently, community-associated (CA) and livestock-associated MRSA clones have also emerged, some of which have become established in hospitals and other healthcare facilities, and sometimes have displaced

previously predominant HA clones. Importantly, MRSA can frequently exhibit resistance to a wide range of clinically relevant antibiotics, which limits treatment options and complicates patient management and outcomes.

Investigating routes of transmission and spread of MRSA in healthcare facilities have conventionally been undertaken by combining available epidemiological information with data from DNA-based typing systems such as pulse-field gel electrophoresis typing, *spa* typing, multilocus sequence typing, and more recently, DNA microarray profiling. However, these approaches can frequently lack the discriminatory ability to differentiate between MRSA isolates in healthcare environments where a relatively small number of clones may predominate. The advent of high-throughput whole-genome sequencing (WGS) over the last decade with the development of affordable, easy-to-use benchtop DNA sequencing platforms, associated sequencing chemistry and bioinformatics tools, has revolutionized studies of MRSA epidemiology and evolution. The significantly enhanced discriminatory power and resolution afforded by WGS has also provided hitherto unimaginable insights into the origins, emergence and factors that drive the evolution of specific MRSA clones. Furthermore, WGS has highlighted the very significant contributions of mobile genetic elements (MGEs) encoding virulence factors and resistance genes from coagulase-negative staphylococcal (CoNS) species to the emergence and evolution of MRSA.

This Research Topic brings together a collection of original research articles and up-to-date reviews that highlight the significant impact WGS is having on our understanding of the epidemiology and routes of transmission of HA- and CA-MRSA in humans, and the phylogenetics and evolution of specific MRSA clones. The Research Topic also highlights the impact that WGS is having on our understanding of antimicrobial resistance in MRSA by acquisition of MGEs, and the role of specific CoNS species in the origins and evolution of particular MGEs that can promote the survival of MRSA following acquisition. Finally, the Research Topic highlights the immense potential impact of WGS technology in surveillance, rapid pathogen detection, identification of virulence factor profiles and antibiotic resistance genotypes, possibly from clinical samples directly.

Table of Contents

Editorial: New Insights and Updates on the Molecular Epidemiology and Antimicrobial Resistance of MRSA in Humans in the Whole-Genome Sequencing Era

David C. Coleman[1]*, Anna C. Shore[1], Richard V. Goering[2] and Stefan Monecke[3,4]

[1] *Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin, Ireland,* [2] *Department of Medical Microbiology and Immunology, School of Medicine, Creighton University, Omaha, NE, United States,* [3] *Leibniz Institute of Photonic Technology (IPHT), InfectoGnostics Research Campus, Jena, Germany,* [4] *Institute for Medical Microbiology and Hygiene, Medical Faculty "Carl Gustav Carus", Technische Universität Dresden, Dresden, Germany*

Editorial on the Research Topic

New Insights and Updates on the Molecular Epidemiology and Antimicrobial Resistance of MRSA in Humans in the Whole-Genome Sequencing Era

Methicillin resistant *Staphylococcus aureus* (MRSA) have been a significant cause of healthcare-associated (HA) infection globally for several decades and more recently, community-associated (CA) infection. Surveillance of MRSA locally, nationally and internationally is critical in monitoring its emergence and spread and informing prevention strategies. Molecular typing has played a key role in tracking the spread of MRSA for many years using a variety of DNA-based methods such as pulsed-field gel electrophoresis typing, *spa* typing and conventional multilocus sequence typing (MLST). However, these approaches are often inadequately discriminatory. The advent of high-throughput whole-genome sequencing (WGS) has revolutionized MRSA research by enabling the complete genomic sequence of isolates to be determined and compared. This has facilitated enhanced outbreak detection and has provided in-depth insights into the evolution of particular MRSA clones and the emergence of new clones.

The objective of this topic was to assemble a range of articles on important aspects of HA- and CA-MRSA in humans, informed by WGS technology. The research topic consists of seven original research articles and two descriptive reviews which focus on up-to-date perspectives in the areas of molecular epidemiology and spread of MRSA and highlights how WGS has revolutionized the epidemiological and phylogenetic investigation of specific MRSA clones in different parts of the world. The topic also highlights the impact of WGS on our understanding of antimicrobial resistance encoded by mobile genetic elements (MGEs) including staphylococcal cassette chromosome *mec* (SCC*mec*) elements and the role of specific coagulase-negative staphylococcal species (CoNS) in the evolution of MGEs, such as the arginine catabolic mobile element (ACME) which can enhance the survival of MRSA following acquisition.

Two of the studies highlight the usefulness and accuracy of WGS, using different analysis methods, for protracted outbreak investigations of two MRSA clones uncommon in the regions in question.

Earls et al. used WGS to investigate the relatedness of CA clonal complex (CC) 88-MRSA isolates from two separate protracted outbreaks (2009–2011 and 2014–2017) in the neonatal intensive care unit of an Irish hospital and their likely geographic origin. Outbreak I included seven *spa* type t186 isolates from seven patients, whereas outbreak II consisted of 15 *spa* type t786 isolates including 13 patient isolates and two isolates from the same healthcare worker (HCW) recovered 2 years apart. All isolates were identified as sequence type (ST) 78-MRSA-IVa and formed a large cluster in a whole-genome MLST-based minimum spanning tree and exhibited 1-71 pairwise allelic differences. Interestingly, in addition to having a different *spa* type, all outbreak II isolates lacked the *hsdS* gene. The two HCW isolates exhibited one allelic difference despite being recovered 2 years apart. Core-genome MLST and sequence-based plasmid analysis revealed the recent shared ancestry of Irish and Australian ST78-MRSA-IVa. The results of this study revealed that the ST78-MRSA-IVa from both protracted outbreaks formed a homogenous population despite differences in *spa* type and the presence/absence of the *hsdS* gene.

Rubin et al. used WGS to investigate a protracted outbreak of ST97-MRSA-IVa over a 4-year period in a surgical ward in a hospital in Denmark involving 23 patients and two HCWs. Eighteen of the patients had been admitted to the surgical ward, one had shared a ward at another hospital with a patient previously admitted to the same surgical ward and two were family members of MRSA-positive patients. Twelve of the patients were diagnosed with ST97-MRSA-IV in the community. The 25 ST97-MRSA-IVa isolates were remarkably homogenous exhibiting a maximum of 50 single nucleotide polymorphisms (SNPs) for isolates recovered over a 4-year period. A subset of the 15 isolates recovered from 13 patients admitted during 2016 and 2017 and the two HCW isolates exhibited a maximum of 15 SNPs. The authors hypothesized that colonized HCWs may have contributed to the maintenance of the protracted outbreak. Both of these studies highlight how WGS can enhance the detection and analysis of prolonged hospital outbreaks of MRSA and that HCWs with an unknown MRSA carriers state might cause sporadic transmission and sustain an outbreak over several years.

Three of the studies undertook large-scale investigations of the evolution of some key successful MRSA lineages over several decades.

van Hal et al. used WGS to examine 459 ST93 *S. aureus* from Australia, New Zealand, Samoa, and Europe to explore the evolution of ST93, its emergence in Australia and subsequent spread to other countries. In Australia, Panton–Valentine leukocidin-positive ST93-MRSA has rapidly become the most common CA-MRSA lineage. Comparisons with other *S. aureus* genomes indicated that ST93 is an early diverging and recombinant lineage, encompassing segments from the ST59/ST121 lineage. Limited genetic diversity was observed within extant ST93 with the most recent common ancestor dated to 1977. The analysis revealed that an epidemic ST93 population arose from a methicillin-susceptible *S. aureus* (MSSA) ancestor in remote Northern Australia, which has a proportionally large population of native Australians with a high prevalence of skin and soft tissue infections. The analysis further revealed that

methicillin resistance was acquired by ST93 three times in these regions with an ST93-MRSA-IVa clade expanding and spreading to Australia's east coast by 2000. Non-sustained introductions of ST93-MRSA-IVa to the United Kingdom were observed, whereas ST93-MRSA-IVa was sustainably transmitted with clonal expansion within the Pacific Islander population in New Zealand who experience similar socio-economic disadvantages as native Australians. This study highlights the significance of human population factors in the emergence of CA-MRSA.

Challagundla et al. undertook the first large-scale phylogenomic study of CC5-MRSA from the Western Hemisphere in order to elucidate the phylogeny of major clones, their place and time of origin and to map the evolution of associated key characteristics. MRSA clones belonging to CC5 are highly diverse and have been a major cause of HA infection in the Western Hemisphere for decades. The study investigated 598 genome sequences, including 409 newly generated sequences, which identified a geographically well-dispersed early branching CC5-Basal clade consisting of MSSA and MRSA. The basal clade gave rise to two major clades in the early 1960's and early 1970's that underwent major expansions in the Western Hemisphere including a CC5-I clade in South America and a CC5-II clade largely in Central and North America. The clade expansions were preceded by convergent acquisitions of resistance to several antibiotic classes and convergent losses of the *sep* gene encoding staphylococcal enterotoxin p. Unique losses of surface proteins were also noted for both clades. The study also revealed that the recombination rate of CC5 was much lower than previously reported for other *S. aureus* lineages. Overall, the results of this study elucidated the relationships between CC5 clades and clones and identified genomic changes for increased antibiotic resistance and decreased virulence associated with CC5 MRSA clade expansions in the Western Hemisphere. These findings suggest that less virulent and more antibiotic resistant CC5-MRSA clones may be better adapted to spread geographically.

Monecke et al. investigated ST239-MRSA-III, a pandemic strain circulating in many countries globally since the 1970s, which has largely been displaced by other MRSA strains in Europe in recent decades. A total of 184 isolates from 11 different countries were characterized using DNA microarrays that profiled an extensive range of typing markers, virulence and antimicrobial agent resistance genes and SCC*mec* types. Thirty additional isolates were subjected to WGS and, together with published WGS data for 215 additional ST239-MRSA-III isolates, were analyzed *in-silico* for comparison with the microarray. This approach assigned the isolates to 39 different SCC*mec* III subtypes, and to three major and several minor clades. One clade comprising isolates and sequences from Turkey, Romania and other Eastern European countries, Russia, Pakistan, and Northern China was characterized by the integration of a transposon into the *nsaB* gene and by the loss of the *fnbB* and *splE* genes. Another clade, harboring *sasX*/*sesI* was found to be widespread in South-East Asia including China/Hong Kong, and in Trinidad & Tobago. A third, related, but *sasX*/*sesI*-negative clade occurs in Latin America but also in Russia and in the Middle East from where it apparently originated. This study

demonstrated that for pandemic ST239-MRSA-III, analysis of genome markers assigned by array hybridization, multiplex PCR, or by WGS can help assigning clinical isolates to these clades or variants and thus help to identify the likely provenance of an isolate.

Aside from providing unparalleled high-resolution for epidemiological, evolutionary and population structure investigations, WGS applications have also enabled the direct comparison of large, distinctive genomic regions associated with pathogenesis or antimicrobial resistance, amongst distinct isolates and species.

McClure et al. applied WGS to the genomic comparison of four closely related ST239 MRSA strains; one avirulent, two moderately virulent and one virulent, in order to gain insights into which areas of the genome might contribute to enhanced pathogenicity. The results revealed that the most virulent strain harbored the complete and intact *spa* and *lpl* genes, both of which contribute to innate immunity and virulence. In the moderately virulent strains, the *spa* gene was present and identical whereas the *lpl* gene was disrupted. Both the *spa* and *lpl* genes were absent in the avirulent strain. The authors also implicated MGEs encoding virulence and antibiotic resistance determinants in strain pathogenicity differences. The MGE arsenal was largest in the virulent strain, although several similar MGEs were also detected in the other three less virulent strains. This study highlights the sensitive interplay that exists between distinct virulence factors in the overall pathogenicity of a strain.

Although first described as a MGE contributing to the growth and survival capability of the USA300-MRSA clone, the high prevalence and extensive diversity of ACME amongst *S. epidermidis* was revealed by O'Connor et al. using WGS. Five main ACME types and further subtypes ranging between 27 and 117 kb were described in oral *S. epidermidis* isolates recovered from the oral rinse, periodontal pockets and subgingival sites of orally healthy individuals with and without dental implants, and in patients with periodontal disease or infected dental implants. This study suggested that ACME may contribute to the survival of *S. epidermidis* in oral environments, likely due to the contribution of the *arc* and *kdp* operons (encoding an arginine deaminase pathway and potassium transporter, respectively) harbored by ACME types I, II, IV, and V. This study also highlighted the role of CoNS in the evolution of *S. aureus* and MRSA, due to its ability to act as a reservoir for MGEs that contribute to strain fitness and survival.

The review by Miragaia outlines the intricate combination of distinct factors that contributed to the emergence, spread and evolution of MRSA, in addition to the application of WGS in deciphering these individual contributory factors. This article describes the pre-existence of native *mec* genes harbored by staphylococcal species predominantly associated with environmental- and animal-sources, the stepwise assembly

and diversification of staphylococcal cassette chromosome incorporating *mec* (SCC*mec*) elements in other CoNS and subsequent transfer to *S. aureus*. The historical context of clinical treatment and farming practices are discussed in relation to the emergence of MRSA as one of the most important worldwide pandemics, originating in hospitals before spreading into the community and livestock.

Lastly, the mini-review by van Belkum and Rochas provides a succinct overview of current laboratory-based and point-of-care testing for MSSA/MRSA detection. In recent years there has been a significant move from culture-based, phenotypic methods toward molecular detection due to the strong correlation between phenotypic resistance and the presence of *mec* genes encoding methicillin resistance. The review highlights the range of approaches, mostly polymerase chain reaction-based detection methods, used for molecular detection of MRSA/MSSA in the laboratory setting or at the point-of-care. The potential diagnostic impact of WGS technology is discussed against a background of diagnostic, surveillance, and infection control parameters and toward enabling rapid pathogen detection and determination of virulence characteristics and antibiotic resistance profiles, potentially directly from clinical specimens.

These articles and reviews emphasize the significant impact that WGS has had on our understanding of the surveillance, epidemiology, and evolution of MRSA and factors that drive the emergence and spread of specific clones. They also reveal the enormous potential impact WGS technology will have on the future detection, prevention, and treatment of MRSA.

AUTHOR CONTRIBUTIONS

DC, AS, RG, and SM conceived the research topic, recruited the contributing authors and edited the topic manuscripts. DC drafted the editorial. AS, RG, and SM reviewed and edited the draft editorial. All authors approved the final editorial.

FUNDING

Intra-Hospital, Inter-Hospital and Intercontinental Spread of ST78 MRSA From Two Neonatal Intensive Care Unit Outbreaks Established Using Whole-Genome Sequencing

Megan R. Earls[1], David C. Coleman[1], Gráinne I. Brennan[2], Tanya Fleming[2], Stefan Monecke[3,4], Peter Slickers[3,4], Ralf Ehricht[3,4] and Anna C. Shore[1]*

[1] Microbiology Research Unit, Dublin Dental University Hospital, Trinity College, University of Dublin, Dublin, Ireland, [2] National MRSA Reference Laboratory, St. James's Hospital, Dublin, Ireland, [3] Abbott (Alere Technologies GmbH), Jena, Germany, [4] InfectGnostics Research Campus, Jena, Germany

From 2009 to 2011 [transmission period (TP) 1] and 2014 to 2017 (TP2), two outbreaks involving community-associated clonal complex (CC) 88-MRSA *spa* types t186 and t786, respectively, occurred in the Neonatal Intensive Care Unit (NICU) of an Irish hospital (H1). This study investigated the relatedness of these isolates, their relationship to other CC88 MRSA from Ireland and their likely geographic origin, using whole-genome sequencing (WGS). All 28 CC88-MRSA isolates identified at the Irish National MRSA Reference Laboratory between 2009 and 2017 were investigated including 20 H1 patient isolates, two H1 isolates recovered from a single healthcare worker (HCW) 2 years apart, three patient isolates from a second hospital (H2) and one patient isolate from each of three different hospitals (H3, H4, and H5). All isolates underwent DNA microarray profiling. Thirteen international isolates with similar microarray profiles to at least one Irish isolate were selected from an extensive global database. All isolates underwent Illumina MiSeq WGS. The majority of Irish isolates (25/28; all H1 isolates, two H2 isolates and the H3 isolate) were identified as ST78-MRSA-IVa and formed a large cluster, exhibiting 1–71 pairwise allelic differences, in a whole-genome MLST-based minimum spanning tree (MST) involving all Irish isolates. A H1/H2, H1/H3, and H1 HCW/patient isolate pair each exhibited one allelic difference. The TP2 isolates were characterised by a different *spa* type and the loss of *hsdS*. The three remaining Irish isolates (from H2, H4, and H5) were identified as ST88-MRSA-IVa and dispersed at the opposite end of the MST, exhibiting 81–211 pairwise allelic differences. Core-genome MLST and sequence-based plasmid analysis revealed the recent shared ancestry of Irish and Australian ST78-MRSA-IVa, and of Irish and French/Egyptian ST88-MRSA-IVa. This study revealed the homogeneity of isolates recovered during two NICU outbreaks (despite *spa* type and *hsdS* carriage variances), HCW involvement in the outbreak transmission chain and the strain's spread

to two other Irish hospitals. The outbreak strain, CC88/ST78-MRSA-IVa, was likely imported from Australia, where it is prevalent. CC88/ST88-MRSA-IVa was also identified in Irish hospitals and was likely imported from Africa, where it is predominant, and/or a country with a large population of African descent.

INTRODUCTION

Staphylococcus aureus is both a prominent pathogen worldwide and an asymptomatic coloniser of the skin and mucosal membranes of humans and animals (Young et al., 2012). The success of *S. aureus* as a disease-causing agent is largely attributable to its ability to acquire and express a wide variety of virulence-associated and antimicrobial resistance genes, many of which can be transferred horizontally between cells on mobile genetic elements (Malachowa and Deleo, 2010). *Staphylococcus aureus* develops into methicillin-resistant *S. aureus* (MRSA) upon acquisition of *mecA* or *mecC*, located on the Staphylococcal Chromosomal Cassette *mec* (SCC*mec*) mobile genetic element, or the plasmid-located *mecB* gene, all of which encode alternate penicillin-binding proteins, conferring resistance to almost all β-lactam antibiotics (Katayama et al., 2000; Shore et al., 2011; Becker et al., 2018). Molecular epidemiological evidence indicates that MRSA have evolved independently from multiple lineages of methicillin-susceptible *S. aureus* in several environments and MRSA clonal groups are therefore often categorised as healthcare-associated (HCA), community-associated (CA) or livestock-associated (LA) (Lindsay, 2010). In recent years, however, it has become increasingly apparent that clonal groups are not limited to the environment in which they initially arose (Bal et al., 2016) and thus, this type of classification currently serves to define a strain's origin.

In-depth surveillance of MRSA, both locally and internationally, is essential in order to identify modes and routes of transmission and, ultimately, to develop effective strategies to prevent or limit transmission. Whole-genome sequence analysis provides optimal typing resolution, thus enabling the accurate determination of isolate relatedness. Over the past five years, both single nucleotide polymorphism (SNP) and allele-based approaches have been utilised during the analysis of bacterial next generation sequencing data. Initially, SNP analysis alone was generally applied to data sets (Price et al., 2013). While this method provides the highest discriminatory power available, a suitable reference genome is not always available and the wide variety of SNP filters for which parameters must be set can impede inter-study comparisons (Schürch et al., 2018). This approach is therefore ideally suited to the comparison of closely related isolates using study-specific reference genomes. In 2014, Leopold et al. built on traditional multilocus sequence typing (MLST) involving just seven loci, by using all 40 finished *S. aureus* genomes available in GenBank as of June 2013 to devise a core-genome (cg)MLST scheme, consisting of 1,861 loci (Leopold et al., 2014). This method provided a standardised

tool for comparing the stable core-genome of *S. aureus* isolates. The cgMLST approach is therefore particularly well-suited to determining the relatedness of isolates recovered over a relatively long period of time and/or from disparate geographic regions, when the environmental conditions to which isolates have been recently exposed and thus, the accessory genome, are of lesser relevance. Recently, extended versions of this cgMLST scheme, which include accessory genome loci, have also been employed (Roisin et al., 2016; Sabat et al., 2017). This approach is typically referred to as whole-genome (wg)MLST and is suited to local outbreak investigations, during which, the comparison of entire genomes may be both appropriate and beneficial. Although there are no definitive cgMLST or wgMLST thresholds for assigning isolate relatedness, it has been suggested that a difference of ≤24 (core genome or whole-genome) alleles may be used as an approximate clonality guideline, however, the longer the time period over which the isolates were recovered, the higher the possibility of their exceeding this threshold (Schürch et al., 2018). Furthermore, sequence-based plasmid analysis may also be used during surveillance or outbreak investigations, although significant bioinformatics expertise is required to optimise short-read plasmid assemblies (Orlek et al., 2017).

Although whole-genome sequencing (WGS) offers optimal typing resolution, conventional molecular typing methods of moderate discriminatory power, such as *S. aureus* protein A (*spa*) typing, are commonly used during hospital outbreak investigations (Frénay et al., 1996). Furthermore, for retrospective compatibility with previous studies, MRSA isolates are still assigned to a MLST clonal complex (CC) and/or sequence type (ST) (Robinson and Enright, 2004) and to one of 13 main SCC*mec* types (Shore and Coleman, 2013; Wu et al., 2015; Baig et al., 2018). Sequence types share a CC if at least five of their seven traditional MLST alleles are identical to at least one other ST in the CC (Feil et al., 2003). The application of such typing techniques also facilitates global MRSA surveillance, allowing for the broader classification MRSA strains and their general categorisation as pandemic, endemic or sporadic. For example, the CA CC88 clones, ST78-MRSA-IV, and ST88-MRSA-IV/V, have both achieved endemic status in particular geographic regions (Monecke et al., 2011). ST78-MRSA-IV is a Panton-Valentine leukocidin (PVL)-negative clone that usually harbours the antimicrobial resistance genes *blaZ* and *erm*(A) (Monecke et al., 2011). First isolated in remote Western Australia in 1995 (O'Brien et al., 2009), ST78-MRSA-IV was the fourth most prevalent clone detected in Australia in 2012, accounting for 3.6% of all MRSA and 5.1% of all CA-MRSA (Coombs et al., 2014). Reports of this clone elsewhere, however, are lacking, with

the exception of a single isolate from Germany (Monecke et al., 2011). ST88 MRSA includes both PVL-positive and negative strains, often harbours the exfoliative toxin gene, *etA*, and is generally associated with SCC*mec* types IV and V, although ST88-MRSA-VI have been identified sporadically in Western Australia (Monecke et al., 2011). Although prevalent in the Far East (Ozaki et al., 2009; Zhang et al., 2009; Qiao et al., 2014) and predominant in Africa, where it accounts for 24.2–83.3% of all MRSA isolates (Breurec et al., 2011; Schaumburg et al., 2011), ST88-MRSA is sporadic in Europe and the Middle East (Monecke et al., 2012; RuŽičková et al., 2012; Vindel et al., 2014). CC88 MRSA have also been described in bulk tank milk and retail food (Parisi et al., 2016; Raji et al., 2016).

MRSA is endemic in Irish hospitals, where the HCA ST22-MRSA-IV strain has predominated since 2002 (Irish National Meticillin-Resistant *Staphylococcus aureus* Reference Laboratory, 2016). An extensive diversity of other MRSA strains and clones has also been identified in Ireland, several of which have been involved in significant hospital outbreaks, including CA strains ST772-MRSA-V, and ST1-MRSA-IV (Brennan et al., 2012; Kinnevey et al., 2014; Earls et al., 2017). Between 2009 and 2011, a suspected protracted outbreak occurred in the Neonatal Intensive Care Unit (NICU) of an Irish hospital (H1), in which seven CC88-associated MRSA-t186 isolates were identified during patient screening. Interestingly, a second suspected protracted outbreak involving a CC88-associated *spa* type occurred in the same NICU between 2014 and 2017, in which 15 MRSA-t786 screening isolates were recovered. Infants in NICUs are particularly vulnerable to serious infection and MRSA colonisation increases their risk of nosocomial infection (Geva et al., 2011). The present study used WGS to achieve three main objectives. Firstly, this study aimed to determine the relationship between isolates from both outbreaks and to identify putative transmission events. Secondly, the relatedness of the outbreak isolates to other CC88 MRSA identified in Ireland was investigated. Finally, considering that CC88 MRSA is not usually associated with Ireland or indeed other European countries, the present study sought to determine the relatedness of all CC88 MRSA recovered in Ireland to international comparator isolates.

MATERIALS AND METHODS

Isolates

All 28 CC88-MRSA isolates identified at the Irish National MRSA Reference Laboratory (NMRSARL) between January 2009 and February 2017 were investigated. Twenty-two of these isolates were recovered in the NICU of H1 during two suspected outbreaks which occurred from 2009 to 2011 (seven isolates from seven inpatients) and 2014 to 2017 (13 isolates from 13 inpatients and two isolates from one healthcare worker [HCW]), respectively. The outbreaks were initially suspected due to the number of isolates identified and their recovery in a single unit in the hospital. While this study included only one isolate per patient, two isolates (W19 and W28) were included from a single HCW, as they were recovered two years apart (in 2015 and 2017, respectively). The remaining six CC88 isolates from Ireland were recovered from inpatients of a second hospital [H2; *n* = 3 (2

NICU and one adult ward)] and in three additional hospitals (H3, *n* = 1; H4, *n* = 1; H5, *n* = 1). All Irish isolates were recovered during colonisation screening. Each isolate is represented by a letter, indicating recovery from either a patient (P) or healthcare worker (W), followed by a number, indicating the order in which the isolates were recovered (**Table 1**). Fifteen CC88-MRSA isolates recovered between 2001 and 2017 in Australia (*n* = 4), France (*n* = 3), Germany (*n* = 5), Tanzania (*n* = 2), and Egypt (*n* = 1) were also included in this study for comparison to the Irish isolates (**Table 1**). These international isolates were selected from the in-house strain collections at Abbott ([Alere Technologies GmbH] Jena, Germany) and the Institute for Medical Microbiology and Hygiene, Technical University of Dresden (Dresden, Germany) based on their genotypic similarity (*n* = 13) or dissimilarity (*n* = 2) to the Irish isolates from the present study (see DNA microarray section below for detailed description of international isolate selection). The strain collections include approximately 22,000 global *S. aureus* isolates, a selection of which have been previously reported (Monecke et al., 2008, 2011, 2016). Each international isolate is represented by a letter, indicating the country of recovery, followed by a number, indicating the order in which the isolates were recovered (**Table 1**).

Isolates were identified as *S. aureus* using the tube coagulase test or the Vitek MS Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry system (Vitek, bioMérieux, Marcy l'Etoile, France) according to the manufacturer's instructions. Methicillin resistance was detected using 30-μg cefoxitin disks (Oxoid Ltd., Basingstoke, United Kingdom) in accordance with European Committee of Antimicrobial Susceptibility Testing methodology and interpretive criteria (European Committee on Antimicrobial Susceptibility Testing, 2017) or using the automated VITEK 1 or VITEK 2 systems (bioMérieux, Nuertingen, Germany). Isolates were stored at −80°C on individual Protect Bacterial Preservation System cryogenic beads (Technical Services Consultants Ltd., Heywood, United Kingdom).

Phenotypic Susceptibility Testing

The susceptibility of all isolates was determined against a panel of 19 antimicrobial agents, in addition to cefoxitin, by disk diffusion using European Committee of Antimicrobial Susceptibility Testing methodology (European Committee on Antimicrobial Susceptibility Testing, 2017), and previously described reference strains and interpretative criteria (McManus et al., 2015). The 19 agents tested were amikacin, ampicillin, chloramphenicol, clindamycin, ciprofloxacin, erythromycin, fusidic acid, gentamicin, kanamycin, linezolid, mupirocin, neomycin, rifampicin, streptomycin, sulphonamide, tetracycline, tobramycin, trimethoprim, and vancomycin.

spa Typing

Genomic DNA for *spa* typing was extracted using the 6% InstaGene matrix solution, according to the manufacturer's instructions (BioRad, München, Germany). The variable X region in the *spa* gene of each isolate underwent PCR amplification using the primers and thermal cycling conditions

TABLE 1 | Epidemiological, phenotypic and genotypic details of the 43 CC88-MRSA isolates investigated.

Source	Isolate/patient no.	Month and/or year of recovery	Sequence type[a]	SCC*mec* type/subtype[b]	*spa* type[c]	Antimicrobial resistance profile[d]	Antimicrobial resistance and virulence-associated genes[e]
Ireland-H1	P1	Jan 2009	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P2	Apr 2009	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P3	Mar 2010	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P4[f]	Nov 2010	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P5[f]	Nov 2010	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P6	Aug 20/11	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	P7	Sep 2011	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, sak, scn, sec, sel*
	P10	Mar 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P11	Mar 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P13	Sep 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P14	Oct 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P15	Nov 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P18	Jan 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	W19[g]	Jan 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P20	June 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P21	Sep 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P22	Oct 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P23	Nov 2015	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P24	Feb 2016	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P26	June 2016	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P27	June 2016	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	W28[g]	Feb 2017	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
Ireland-H2	P8	Aug 2013	88	IVa- CMFT503	t786	Ap, Cm, Fx, Tp	*blaZ, cadX, cat, dfrS1, mecA, hsdS, lukX, sak, chp, scn*
	P16	Nov 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
	P17	Dec 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
Ireland-H3	P12[h]	Sep 2014	78	IVa-MW2	t786	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, lukX, sak, scn, sec, sel*
Ireland-H4	P9	Sep 2013	88	IVa- CMFT503	t786	Ap, Fx, Tp	*blaZ, cadX, dfrS1, mecA, etA, hsdS, lukX, sak, chp, scn*
Ireland-H5	P25	Mar 2016	88	IVa- CMFT503	t786	Ap, Fx, Tp	*blaZ, cadX, dfrS1, mecA, etA, hsdS, lukX, sak, chp, scn*
Australia	A1	2001	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn*
	A2	2002	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	A3	2008	78	IVa-MW2	t186	Ap, Er, Fx, Tp	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
	A4	2008	78	IVa-MW2	t186	Ap, Fx, Tp	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn, sec, sel*
Egypt	E1[i]	2014	88	IVa- CMFT503	t13712	Ap, Fx, Tp	*blaZ cadX, dfrS1, mecA, hsdS, lukX, sak, chp, scn*
France	F1	2002	88	IVa- CMFT503	t186	Ap, Er, Fx, Tp	*blaZ, cadX, dfrS1, erm*(C), *mecA, etA, hsdS, lukX, sak, chp, scn*
	F2	2002	88	IVa- CMFT503	t786	Ap, Er, Fx, Tp	*blaZ, cadX, dfrS1, erm*(C), *mecA, etA, hsdS, lukX, sak, chp, scn*
	F3	2002	88	IVa- CMFT503	t690	Ap, Fx, Te, Tp	*blaZ, cadX, dfrS1, mecA, tet*(K), *vga*(A), *etA, hsdS, lukX, sak, chp, scn*
Germany	G1	2008	78	IVa-MW2	t186	Ap, Er, Fx	*blaZ, cadX, erm*(A), *mecA, hsdS, lukX, sak, scn*
	G2	2016	88	IVa- CMFT503	t1028	Ap, Fx, Te, Tp	*blaZ, cadX, dfrS1, mecA, tet*(K), *hsdS, lukX, sak, chp, scn*
	G3	2017	88	IVa- CMFT503	t786	Ap, Fx, Tp	*blaZ, cadX, dfrS1, mecA, hsdS, lukX, sak, chp, scn*
	R1	2014	88	IVa-MW2	t17863	Ap, Er, Fx	*blaZ, cadX, erm*(C), *mecA, hsdS, lukF, lukS, lukX, sak, chp, scn*

(Continued)

TABLE 1 | Continued

Source	Isolate/patient no.	Month and/or year of recovery	Sequence type[a]	SCC*mec* type/subtype[b]	*spa* type[c]	Antimicrobial resistance profile[d]	Antimicrobial resistance and virulence-associated genes[e]
	R2	2017	88	IVa-MW2	t5041	Ap, Fx, Te, Cp	*blaZ, cadX, mecA, tet*(K), *hsdS, lukF, lukS, lukX, sak, chp, scn*
Tanzania	T1	2016	88	IVa- CMFT503	t690	Ap, Fx, Tp	*blaZ, cadX, dfrS1, mecA, hsdS, lukX, sak, chp, scn*
	T2	2016	88	IVa- CMFT503	t690	Ap, Fx, Tp	*blaZ, cadX, dfrS1, mecA, vga*(A), *etA, hsdS, lukX, sak, chp, scn*

[a]*Sequence types (STs) were assigned using Ridom SeqSphere+ version 4.1 (Ridom GmbH, Germany). Allelic profiles: ST78, 22-1-14-23-12-53-31; ST88, 22-1-14-23-12-4-31.*

[b]*All SCCmec subtypes were detected using an SCCmec subtyping DNA microarray (Monecke et al., 2016). All Irish isolates underwent in silico analysis for predicted DNA SCCmec subtype microarray hybridisation profiles, while all other isolates underwent real-life DNA microarray analysis. Both SCCmec subtypes IVa-MW2 (GenBank accession: BA000033.2) and IVa-CMFT503 (GenBank accession: HF569113.1) have been described previously (Monecke et al., 2016).*

[c]*All H1 t186 isolates were involved in an outbreak between 2009 and 2011. All H1 t786 isolates were involved in an outbreak between 2014 and 2017. spa repeat successions: t186, 07-12-21-17-13-13-34-34-33-34; t786, 07-12-21-17-13-34-34-33-34; t690, 07-12-21-17-13-13-34-34-33-34; t1028, 07-34-33-34; t5041, 07-12-21-17-13-13-34-34-34-34-33-34; t13712, 07-12-21-17-13; t17863, 07-12-12-13-13-13-34-33-34.*

[d]*Antimicrobial resistance phenotypes were determined by testing the susceptibility of isolates to a panel of 20 antimicrobial agents including amikacin, ampicillin (Ap), cefoxitin (Fx), chloramphenicol (Cm), ciprofloxacin (Cp), clindamycin, erythromycin (Er), fusidic acid, gentamicin, kanamycin, linezolid, mupirocin, neomycin, rifampicin, streptomycin, sulphonamide, tetracycline, tobramycin, trimethoprim (Tp) and vancomycin.*

[e]*All antimicrobial resistance and virulence-associated genes were detected using the S. aureus Genotyping Kit 2.0 system [Abbott (Alere Technologies GmbH), Jena, Germany]. All Irish isolates underwent in-silico analysis for predicted Genotyping Kit 2.0 DNA microarray hybridisation profiles, while all other isolates underwent real-life DNA microarray analysis.*

[f]*Isolates P4 and P5 were recovered from twins on the same day.*

[g]*Isolates W19 and W28 were recovered from the same healthcare worker two years apart.*

[h]*The patient from whom isolate P12 was recovered had been transferred from H1.*

[i]*Isolate E1 was recovered from a buffalo. All other isolates were recovered from humans.*

H, hospital; NA, not available.

described by the European Network of Laboratories for Sequence Based Typing of Microbial Pathogens (SeqNet; http://www.seqnet.org). Resulting PCR products were purified using the GenElute PCR clean-up kit (Sigma-Aldrich Ireland Ltd., Wicklow, Ireland) and were sequenced commercially (Source Bioscience, Waterford, Ireland) using an ABI 3730xl Sanger sequencing platform. The Ridom StaphType software package version 1.5 (Ridom Gmbh, Würzburg, Germany) was used for *spa* sequence analysis and *spa* type assignment.

Genotyping and SCC*mec* Subtyping Using DNA Microarrays

The 15 international CC88 isolates underwent DNA microarray profiling using the *S. aureus* Genotyping Kit 2.0 system [Abbott (Alere Technologies GmbH)] and an additional SCC*mec* typing array (Monecke et al., 2016). The *S. aureus* Genotyping Kit 2.0 DNA microarray detects 333 target sequences, corresponding to approximately 170 different genes and their allelic variants, and encoding antimicrobial resistance and virulence-associated genes, species, and typing markers, and several SCC-associated marker genes. It assigns *S. aureus* isolates to MLST CCs/STs and SCC*mec* types. Detailed descriptions of the relevant genes, primers, and probes have been previously described (Monecke et al., 2008). The SCC*mec* typing array targets an additional 83 distinct sequences that are variably present in the SCC*mec* element and which form the basis of a previously described system that distinguishes between 61 different SCC*mec* subtypes (Monecke et al., 2016). Detailed descriptions of these SCC*mec*-linked genes/alleles and their corresponding primers and probes have been previously published (Monecke et al., 2016). Genomic DNA was extracted for both DNA microarray profiling and SCC*mec* array subtyping by enzymatic lysis using the *S. aureus*

Genotyping Kit 2.0 [Abbott (Alere Technologies GmbH)] and the Qiagen DNeasy blood and tissue kit (Qiagen, West Sussex, United Kingdom). As all Irish isolates underwent WGS as part of the present study (see below), their genome sequences underwent *in silico S. aureus* Genotyping Kit 2.0 microarray profiling and SCC*mec* array subtyping. Virtual DNA array hybridisation patterns were generated whereby contigs were searched for probe binding sites and signal strength was dictated by the number of nucleotide mismatches, as previously described (Monecke et al., 2016). The CC, DNA microarray profile and SCC*mec* type/subtype was determined for each Irish isolate. Thirteen international isolates, with a similar DNA microarray profiles to at least one Irish isolate, were selected from the aforementioned global database, for WGS (**Table 1**). Two international "reference isolates" with dissimilar DNA microarray profiles to any of the Irish isolates were also selected for WGS, as controls (**Table 1**).

Whole-Genome Sequencing

All isolates underwent WGS using genomic DNA extracted as described above for DNA microarray profiling. DNA quality was assessed by UV absorbance using the NanoDrop spectrophotometer 2000 (ThermoFisher Scientific, Dublin, Ireland) and dilutions were performed using the Qubit Fluorometer 3.0 (ThermoFisher Scientific). The Nextera XT DNA Library Preparation Kit (Illumina, Eindhoven, The Netherlands) was used according to the manufacturer's instructions and libraries underwent paired-end sequencing using the 500-cycle MiSeq Reagent Kit v2 (Illumina). Libraries were scaled to exhibit at least 100x coverage and the quality of each sequencing run was assured following cluster density and Q30 assessment.

Whole-Genome Sequence Analysis
wgMLST and cgMLST

WGS data were analysed using the wgMLST scheme available in BioNumerics v7.6 (Applied Maths, Sint-Martens-Latem, Belgium) consisting of 3,904 *S. aureus* wgMLST loci (Roisin et al., 2016), including 1,861 cgMLST loci (Leopold et al., 2014). In order to ensure that all relevant alleles present were detected, two separate algorithms were used to generate a consensus whole-genome MLST profile for each isolate. The first method determined locus presence/absence and allelic identity using an assembly-free k-mer approach. The second, assembly-based method, used a BLAST approach to detect alleles on contigs assembled using the SPAdes software v3.7.1 (Bankevich et al., 2012) incorporated into BioNumerics. Default base correction parameters were applied and all contigs below 1000 bp were removed. The default settings were used for both the assembly-free and assembly-based algorithms. The quality of the sequence read sets, *de novo* assemblies, and assembly-free, and assembly-based allele calls, were assessed using the quality statistics window in BioNumerics and are detailed in Dataset S1. Traditional MLST sequence types were assigned using Ridom SeqSphere+ version 4.1 (Ridom GmbH, Germany).

SNP Analysis

Isolates confirmed to be closely related following wgMLST subsequently underwent SNP analysis using a study-specific reference sequence. The SPAdes assembly of isolate P6 was chosen as the reference sequence due to both its central position in the wgMLST-based minimum spanning tree (MST) cluster and the high quality of its assembly. The BioNumerics Power Assembler mapping algorithm was used to create a consensus sequence for each sample and a pairwise distance matrix was generated. SNPs were called exclusively in positions shared by all samples. Only SNPs with at least 40x coverage were considered. Potentially indel-related SNPs, occurring within 12 bp of each other, were removed. Positions with ambiguous base calls and SNPs in repetitive regions were excluded.

Minimum Spanning Trees

Minimum spanning trees were constructed firstly, involving the Irish isolates exclusively and secondly, involving both the Irish and international isolates. For the Irish isolates, in order to identify the most appropriate analysis method, three separate MSTs were generated based on cgMLST, wgMLST or SNP data, and were examined in tandem with all available epidemiological and genotypic information. As the Irish and international isolates were recovered over 16 years and from disparate geographic regions, the construction of a cgMLST-based MST was deemed appropriate in this instance. All MSTs were generated using the permutation resampling function and default priority rule set in BioNumerics. The resampling support for each branch was examined to ensure the validity of the general MST structure.

Sequence-Based Plasmid Analysis

All isolate genomes underwent sequence-based plasmid analysis. Sequence read sets were assembled using SPAdes v3.11.1 (Bankevich et al., 2012) with a final kmer size of 127. All

contigs under 500 bp and all conitgs with kmer coverage less than 3.0 were excluded. For each isolate, a scatter plot was generated depicting the GC content versus coverage for each contig. Putative plasmid-derived contigs were differentiated from chromosomal-derived contigs based on their elevated coverage, low GC content and the identity of their first and last 127 nucleotides, indicative of a circular replicon. All putative plasmid-derived contigs were blasted against The National Centre for Biotechnology Information database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and those which mapped to a known plasmid sequence in GenBank were considered to be confirmed plasmids. Any plasmid types present in the Irish isolates or in both the Irish and international isolates were identified. For each such plasmid type, a multi-sequence global alignment was constructed including all the newly identified plasmid sequences and the GenBank reference sequence, using MAFT v7.273 (Katoh et al., 2002).

RESULTS
Two Distinct Clusters of CC88-MRSA Isolates

The majority of Irish isolates were identified as *spa* type t786 ($n = 21$), while those remaining were identified as t186 ($n = 7$; **Table 1**). All seven t186 isolates were recovered during the first suspected H1 outbreak, while the t786 isolates were recovered either during the second suspected H1 outbreak ($n = 13$) or from one of the four alternative hospitals (H2-H5; **Table 1**). The majority of Irish isolates (25/28) were also identified as ST78-MRSA-IVa, harbouring an SCC*mec* type IVa element corresponding to that identified in the MW2 MRSA strain (GenBank accession: BA000033.2). This included all t186 isolates (7/7) and 18/21 t786 isolates. The remaining three t786 isolates were assigned to ST88, a single locus variant of ST78, and harboured an SCC*mec* type IVa element corresponding to that identified in the CMFT503 MRSA strain (GenBank accession: HF569113.1; **Table 1**). The presence of a hypothetical SCC*mec* terminus protein, Q9XB68, in the MW2-like SCC*mec* element and the presence of both an alternate SCC*mec* terminus protein, SCC*mec* terminus 01, and the LytTR domain DNA-binding regulator, Q931B7, in the CMFT503-like SCC*mec* element, distinguish the two SCC*mec* type IVa elements (Monecke et al., 2016).

Interestingly, a highly conserved gene within *S. aureus* lineages, the type I restriction modification system gene, *hsdS* (Waldron and Lindsay, 2006), was absent from all t786 ST78-MRSA-IVa isolates. It was subsequently noted that, according to the cgMLST-based MST, this unusual deletion had occurred twice within a relatively small population (Figure S1). Importantly, however, the wgMLST-based MST (**Figure 1**) indicated that the *hsdS* deletion occurred just once within this population. It was therefore concluded that the wgMLST-based tree likely depicted the evolutionary path of this strain more accurately than the cgMLST tree. To confirm/dispute this finding, a SNP-based MST was generated involving the relevant isolates (Figure S2). The structure of this tree was in agreement with that of

the wgMLST MST, confirming that the *hsdS* deletion likely occurred once during the strain's spread. Ultimately, considering that the application of SNP analysis was not appropriate for all 28 Irish isolates, the wgMLST-based MST was selected for detailed data interpretation. Therefore, any allelic distances stated herein between Irish isolates exclusively, refer to wgMLST loci, while those stated between Irish and international isolates, or between international isolates exclusively, refer to cgMLST loci, as outlined in the methods.

Whole-genome MLST further supported the differentiation of isolates as suggested by their SCC*mec* subtypes and traditional STs, grouping all 25 ST78-MRSA-IVa isolates into a large cluster at one end of a MST, while the three ST88-MRSA-IVa isolates dispersed at the opposite end of the tree. The ST78-MRSA-IVa and ST88-MRSA-IVa isolates differed by a minimum of 325 alleles and exhibited average pairwise allelic distances of 23.8 (min. = 1; max. = 71) and 140.7 (min. = 81; max. = 211), respectively (**Figure 1**).

ST78-MRSA-IVa

The 25 ST78-MRSA-IVa isolates which grouped into the large wgMLST-MST cluster were recovered at intervals of 0–30 months and included all H1 isolates, two H2 isolates (P16 and P17) and the H3 isolate (P12). All patients from whom ST78-MRSA-IVa isolates were recovered were neonates. The patient from whom the H3 isolate was recovered, had recently been transferred from H1. All ST78-MRSA-IVa isolates exhibited resistance to both ampicillin and erythromycin, and harboured the β-lactamase resistance gene, *blaZ*, the macrolide-lincosamide, and streptogramin B resistance gene, *erm*(A), the cadmium tolerance gene, *cadX*, the immune evasion complex (IEC) genes *sak,* and *scn* (IEC type E) and the enterotoxin genes *sec,* and *sel* (**Table 1**). Isolate P7 was the only isolate that lacked the leukocidin homologue, *lukX*. The maximum distance observed between any two directly linked nodes was 32 alleles, detected between t186 isolates, P1 and P7, which were recovered almost 3 years apart during suspected outbreak 1. All other directly linked isolates exhibited 1–19 allelic differences, significantly fewer than the recently proposed approximate clonality threshold of 24 alleles. This indicated a high degree of relatedness between the vast majority of directly linked isolates and a significant relationship between all isolates within the cluster network (**Figure 1**). Furthermore, there were no apparent sub-clusters dictated by *spa* type, suggesting that the "two outbreak strains" were homogeneous. Specifically, the branch that linked isolates P6 (t186) and P10 (t786) constituted the only direct link between the t786 and t186 isolates. However, this branch represented an allelic distance of 18, lower than both those of 19 and 32, each of which was observed elsewhere in the MST cluster. Interestingly, while the largely linear structure of the t186 isolates indicated that the outbreak strain was transmitted in a relatively sequential manner between 2009 and 2011, the highly branched network of t786 isolates suggested that a more complex transmission chain was established between 2014 and 2017 (**Figure 1**).

Isolates P4 and P5, which exhibited eight allelic differences, were recovered from twins on the same day, suggesting that
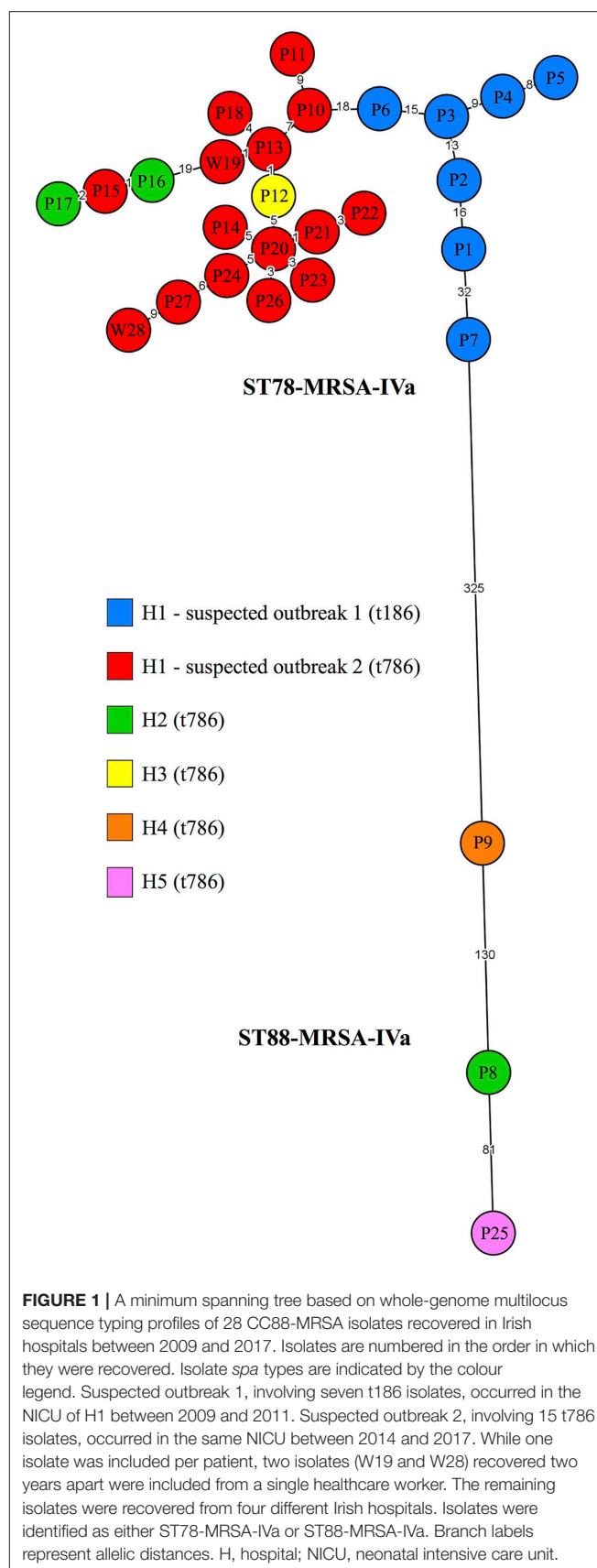
FIGURE 1 | A minimum spanning tree based on whole-genome multilocus sequence typing profiles of 28 CC88-MRSA isolates recovered in Irish hospitals between 2009 and 2017. Isolates are numbered in the order in which they were recovered. Isolate *spa* types are indicated by the colour legend. Suspected outbreak 1, involving seven t186 isolates, occurred in the NICU of H1 between 2009 and 2011. Suspected outbreak 2, involving 15 t786 isolates, occurred in the same NICU between 2014 and 2017. While one isolate was included per patient, two isolates (W19 and W28) recovered two years apart were included from a single healthcare worker. The remaining isolates were recovered from four different Irish hospitals. Isolates were identified as either ST78-MRSA-IVa or ST88-MRSA-IVa. Branch labels represent allelic distances. H, hospital; NICU, neonatal intensive care unit.

parallel or sequential acquisition may have occurred in this instance (**Figure 1**). One of the H2 isolates, P16, and a H1 isolate (P15) recovered 6 days before isolate P16, exhibited one allelic difference, strongly indicating that the outbreak strain spread between these two hospitals and suggesting transmission from the same source (**Figure 1**). Isolate P16 and the second H2 isolate (P17), which was recovered 44 days after isolate P16, exhibited three allelic differences, indicating further spread of this strain in H2 (**Figure 1**). Similarly, the only H3 isolate (P12) and a H1 isolate (P13) recovered 4 days after the H3 isolate, exhibited one allelic difference, clearly indicating that the outbreak strain spread from H1 to H3, and suggesting transmission from the same source (**Figure 1**). Interestingly, the two t786 isolates (W19 and W28) recovered from the same HCW two years apart exhibited 20 allelic differences, indicating that the strain had either altered over time *in vivo*, or that the HCW transiently carried different variants of the strain. Isolates W19 and W28 differed from the other t786 isolates by 1–21 (average:10.5) and 9–37 (average: 21.3) alleles, respectively, suggesting transmission of the outbreak strain between patients and the HCW (**Figure 1**).

ST88-MRSA-IVa

The three ST88-MRSA-IVa isolates, which exhibited an average of 140.7 pairwise allelic differences, included the final H2 isolate (P8), and the H4 (P9), and H5 (P25) isolates, all of which were t786 (**Table 1** and **Figure 1**). None of the patients from whom ST88-MRSA-IVa isolates were recovered were neonates (patients were aged 15 months, 25 years and 66 years). Two of the ST88-MRSA isolates (P8 and P9) were recovered from patients with names suggestive of a family connection to an African country. The phenotypic resistance profiles varied slightly amongst the ST88-MRSA-IVa isolates, all of which exhibited resistance to both ampicillin and trimethoprim, while isolate P8 exhibited chloramphenicol resistance (**Table 1**). The ST88-MRSA-IVa isolates also exhibited slightly differing genotypic profiles, all harbouring resistance genes *dfrSI*, encoding trimethoprim resistance, *blaZ,* and *cadX*, and the IEC genes *chp*, *sak,* and *scn* (IEC type B), while isolate P8 carried the chloramphenicol resistance gene, *cat*, and isolates P9 and P25 harboured *etA*, encoding exfoliative toxin A (**Table 1**). Considering these differences, the lack of epidemiological links and most importantly, the number of alleles by which they differed, these three isolates did not appear to be closely related.

Relatedness of Irish and International CC88 MRSA

Five of the 13 international CC88-MRSA isolates exhibiting similar array profiles to the Irish isolates were identified as ST78-MRSA-IVa-MW2. This included one German isolate (G1) recovered in 2008 and four Australian isolates, A1-A4, recovered in 2001, 2002, 2008 and 2008, respectively (**Table 1**). The remaining eight international isolates exhibiting similar array profiles to the Irish isolates were identified as ST88-MRSA-IVa-CMFT503. This included three French isolates (F1-3) recovered in 2002, two Tanzanian isolates (T1 and T2) recovered in 2016, one Egyptian isolate recovered in 2014

(E1) and two German isolates, G4 and G5, recovered in 2016 and 2017, respectively (**Table 1**). The two international reference isolates (R1 and R2 recovered in Germany in 2014 and 2017, respectively), were identified as ST88-MRSA-IVa-MW2. Following the construction of a cgMLST-based MST including all Irish and international isolates (Figure S3), international isolates R1 and R2 were excluded from further analysis as they failed to cluster with any other isolates, differing from their (shared) most closely related isolate (A1) by 189 and 198 alleles, respectively.

Irish and International ST78-MRSA-IVa Isolates

The five international ST78-MRSA-IVa isolates harboured the same resistance and virulence-associated genes as the Irish ST78-MRSA-IVa isolates (**Table 1**). The Irish and international ST78-MRSA-IVa isolates also exhibited very similar phenotypic susceptibility profiles, with both isolate groups exhibiting ampicillin and erythromycin resistance, while two international isolates (A3 and A4) also exhibited trimethoprim resistance. Interestingly, all international ST78-MRSA-IVa isolates were identified as t186 and none exhibited the *hsdS* deletion that characterised the Irish t786 ST78-MRSA-IVa isolates. Following the generation of a cgMLST-based MST including Irish and international isolates (**Figure 2**), the international ST78-MRSA-IVa isolates grouped in relatively close proximity to the Irish ST78-MRSA-IVa cluster and exhibited an average pairwise allelic distance of 68.7 (min. = 36; max. = 105; **Figure 2**). Specifically, isolates A2-A4 and G1 all radiated independently from isolate A1, which was the only isolate that linked directly to the t186 side of the Irish cluster (**Figure 2**). Isolate A1 (recovered in 2001 in Australia) and its most closely related Irish isolate (P1; recovered in 2009) exhibited 60 allelic differences which, considering the disparate geographic regions and different time periods in which they were recovered, suggested a significant degree of relatedness between these two isolates.

All Irish and international ST78-MRSA-IVa isolates harboured a 21 kb plasmid encoding *blaZ* and *cadX*, corresponding to plasmid pWBG763 (GenBank accession number: GQ900467.1). Interestingly, however, the Irish ST78-MRSA-IVa isolates were characterised by a 100 bp deletion in this plasmid. All Irish and two Australian (A2 and A3) ST78-MRSA-IVa isolates also harboured a cryptic 2 kb plasmid corresponding to pWBG764 (GenBank accession number: GQ900468.1). Both plasmids pWBG763 and pWBG764 were originally sequenced from the same ST78-MRSA-IVa strain (WBG8366), which was recovered in remote Western-Australia in 1995.

Irish and International ST88-MRSA-IVa Isolates

The eight international ST88-MRSA-IVa isolates harboured similar resistance and virulence-associated genes to the Irish ST88-MRSA-IVa, all carrying *blaZ*, *cadX*, *dfrS1,* and the IEC genes *chp*, *sak* and *scn* (IEC type B), while exhibiting variable *etA* carriage (**Table 1**). However, two of the international ST88-MRSA-IVa isolates (F1 and F2) also harboured *erm*(C), encoding macrolide resistance, while two others (F3 and G2) harboured *tet*(K), encoding tetracycline resistance and a third pair (F3 and
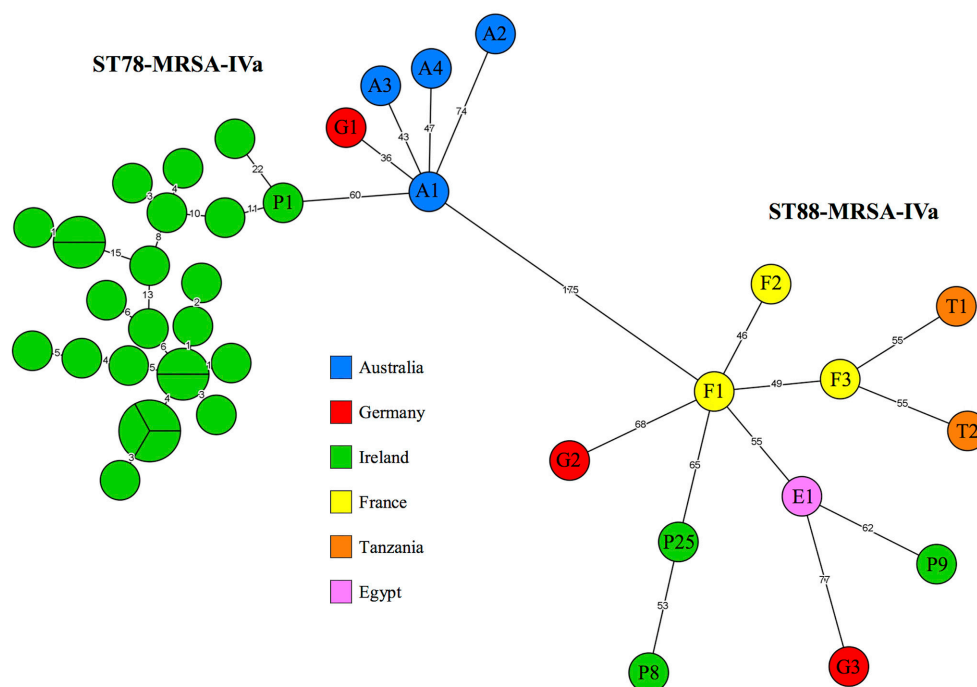
FIGURE 2 | A minimum spanning tree based on core-genome multilocus sequence typing profiles of 28 Irish and 13 international CC88-MRSA isolates recovered between 2001 and 2017. The countries in which the isolates were recovered are indicated by the colour legend. Isolates were identified as either ST78-MRSA-IVa or ST88-MRSA-IVa. Branch labels represent allelic distances.

T2) harboured *vga*(A), encoding streptogramin A resistance, none of which were present in the Irish isolates (**Table 1**). The Irish and international ST88-MRSA-IVa isolates also exhibited similar phenotypic susceptibility profiles, with both isolate groups exhibiting ampicillin and trimethoprim resistance, while two international ST88-MRSA-IVa isolates (F1 and F2) exhibited erythromycin resistance and two others (F3 and G2) exhibited tetracycline resistance, neither of which were observed in the Irish ST88-MRSA-IVa isolates. None of the international ST88-MRSA-IVa isolates exhibited chloramphenicol resistance, which was observed in one Irish ST88-MRSA-IVa isolate (P8). Isolate F1 was identified as t186, while the remaining international ST88-MRSA-IVa isolates were identified as t13712 ($n = 1$), t1786 ($n = 2$), t690 ($n = 3$) or t1028 ($n = 1$; **Table 1**). The eight international ST88-MRSA-IVa isolates formed a dispersed cluster with the three Irish ST88-MRSA-IVa isolates, in which an average pairwise allelic distance of 78.6 was observed (min. = 46; max. = 114; **Figure 2**). Irish isolates P25 and P9 differed from their most closely related international isolates (F1 and E1, respectively) by 65 and 62 alleles, respectively, indicative of shared ancestral genotypes. Irish isolate P8 and its most closely related international isolate (F1) exhibited 188 allelic differences, suggesting a lack of relatedness between the two isolates (**Figure 2**).

All Irish and international ST88-MRSA-IVa isolates harboured a 25 kb plasmid encoding *blaZ* and *cadX*, which corresponded to contig 10 (GenBank accession number: FMNJ01000010.1) of a previously published WGS project

involving an MRSA isolate (GenBank accession number: FMNJ01000000.1) recovered in Tanzania in 2008. Irish ST88-MRSA-IVa isolate P8 also harboured a *cat*-encoding plasmid, which corresponded to contig 32 (GenBank accession number: LFNS01000032.1) of a previously published WGS project involving an ST3019 *S. aureus* isolate recovered in Ghana in 2013 (GenBank accession number: LFNS00000000.1).

DISCUSSION

The present study revealed the homogeneity of isolates involved in two outbreaks in the NICU of an Irish hospital. Although isolate *spa* types and recovery dates suggested that two different CC88-MRSA strains may have been involved in the outbreaks in this NICU, wgMLST revealed that these outbreaks were caused by the same CC88/ST78-MRSA-IVa strain, which spread within the ward during two separate transmission periods (transmission periods 1 and 2). This investigation highlighted both the involvement of a HCW in the outbreak transmission chain and the strain's spread to two other Irish hospitals. A cgMLST-based comparison with international comparator isolates revealed that the outbreak strain was most likely imported from Australia, where it is among the prevalent MRSA clones. This study also identified a second CC88-MRSA clone present in Irish hospitals, ST88-MRSA-IVa, which was likely imported from Africa, where it is predominant, and/or a country with a large population of African ethnic origin.

Transmission period 1 (TP1) involved the intermittent acquisition of the outbreak strain by seven NICU patients over a 32-month period. Interestingly, the topology of the MST indicated that all TP1 isolates, apart from P7, were acquired in a relatively sequential chain of transmission. While the topological characteristics of a phylogenetic tree can reveal invaluable relatedness and transmission details, both previously published studies and epidemiological data must also be drawn upon in order to gain meaningful insights into the dynamics of an outbreak. Notably, previous studies have indicated that patient-to-patient transmission is rare in adult intensive care units (Price et al., 2014; Wesley Long et al., 2014), a finding which is likely applicable to NICUs given the dependency of neonates on adults for mobility. Furthermore, it is unlikely that patients P1-P6 had overlapping stays (excluding twins, P4, and P5), considering the dates on which their isolates were recovered (Table 1). It may therefore be concluded that patient-to-patient transmission did not play a significant role in the outbreak during this period. Similarly, given that previous research suggests that S. aureus can survive a maximum of 90 days on hospital plastics and fabrics (Neely and Maley, 2000), and isolates P1-P6 were recovered at intervals of 4–11 months (Table 1), it is perhaps unlikely that patients P1-P6 acquired the outbreak strain directly from their environment without the involvement of another intermediary factor(s). Finally, previous studies have identified a role for HCWs in the transmission of MRSA to NICU patients (Geva et al., 2011; Brennan et al., 2012; Azarian et al., 2016). Considering these points, it appears highly likely that HCWs were involved in the spread of the outbreak strain during TP1, however, as routine HCW screening did not occur, their exact role cannot be definitively determined. Furthermore, these considerations, in combination with the topology of the MST, indicate that more than one vector was involved in spreading the outbreak strain during TP1. The data suggest two possible scenarios. Firstly, it is possible that a HCW constituted the primary outbreak source, originally seeding the outbreak strain in 2009, transmitting it to patient P1, and again in 2011, transmitting it to patient P7, while a different HCW initiated the strain's spread to patients P2-P6 (Table 1; Figure 1). Alternatively, the data suggest that patient P1 constituted the primary outbreak source and, while one HCW initiated transmission to patients P2-P6, a different HCW, who had also acquired the strain in 2009, eventually transmitted it to patient P7.

In Ireland, neonates are generally screened for MRSA upon admission into a NICU and weekly, thereafter (Irish Department of Health, 2013). Interestingly, however, the outbreak strain identified here was not detected between 2011 and 2014. It is unknown whether any staff changes or staff decolonisation occurred in H1 during this intervening period. Upon reappearing in 2014, the outbreak strain had undergone slight modifications which were detectable using conventional molecular epidemiological typing. Specifically, spa typing indicated that the spa gene had evolved from t186 to t786 (the latter of which is distinguishable from the former by the absence of one repeat unit), while DNA microarray profiling revealed an unusual hsdS deletion. It is highly likely that these alterations occurred locally, either while the strain resided in-vivo in a H1 HCW or during the strain's spread in the community, prior to reintroduction into the NICU.

Transmission period 2 (TP2) involved the acquisition of the outbreak strain, ST78-MRSA-IVa, by 20 patients and one HCW, over a 35-month period. Interestingly, the MST indicated that the vector from which patient P6 acquired the outbreak strain (during TP1), may have constituted the source of the outbreak at the beginning of TP2 (Figure 1). Furthermore, as observed during TP1, it appeared that more than one vector was involved in the spread of the outbreak strain during TP2. This was evident from the significant extension of three TP2 isolates (P15, P16, and P17, recovered in H1, H2, and H2, respectively) from the main body of the cluster in which isolates with both earlier and later recovery dates resided (Figure 1). In contrast to TP1, however, the TP2 isolates were recovered at intervals of 0–7 months suggesting that some TP2 patients may have had overlapping stays (Table 1). This circumstance may have contributed to the establishment of a more complex transmission chain during this time period. Although HCW screening is not mandatory in Ireland, it is indicated if transmission continues on a unit despite active control measures, if epidemiological aspects of an outbreak or strain are unusual, or if they suggest persistent MRSA carriage by staff (Irish Department of Health, 2013). This was likely the basis upon which HCW screening took place during TP2, the extent of which, is unknown. Importantly, an average pairwise distance of 10.5 (range: 1–21) between a H1 HCW isolate (W19) and the remaining TP2 isolates indicated that this HCW was likely directly involved in transmitting the outbreak strain to patients during this period. Similarly, a difference of one allele between both a H1/H2 (P15 and P16) and H1/H3 (P13 and P12) isolate pair, indicated that the outbreak strain spread to two additional hospitals. In the case of H3, it is highly likely that patient P12 acquired the outbreak strain in H1, before being transferred to H3. Similarly, although no known patient transfers occurred between H1 and H2, it is possible that a carrier who was not represented in the present study (i.e. a patient not screened during routine surveillance) was transferred from H1 to H2, during TP2. This is particularly feasible given the high frequency with which patients are transferred between Irish hospitals. However, as the employment of specialist healthcare staff by different hospitals is not uncommon in Ireland, it remains possible that the movement of staff facilitated the inter-hospital spread of this strain.

Genotypic data from both the present investigation and previously published studies were considered while determining the putative geographic origin(s) of the outbreak strain. Firstly, cgMLST indicated that the outbreak strain shared an ancestral genotype with an isolate recovered in Australia (Figure 2). Furthermore, a 21 kb blaZ, and cadX-encoding plasmid was detected in all Australian and Irish ST78-MRSA-IVa isolates, while a second cryptic 2 kb plasmid was detected in two Australian and all Irish ST78-MRSA-IVa isolates. Moreover, both of these plasmids were previously sequenced from the same Australian ST78-MRSA-IVa strain. Finally, ST78-MRSA-IVa is generally reported exclusively from Australia and the rate of travel between Australia and Ireland was consistently high in

the years preceding the study period (Australian Government Department of Immigration Citizenship, 2011). Considering these points, it is highly likely that the outbreak strain was imported from Australia, where it is commonly known as Western Australia MRSA-2 (Coombs et al., 2012). Interestingly, a 2012 study reported that while ST78-MRSA-IVa was the second most prevalent strain among HCWs in a Western Australian hospital, it was associated exclusively with persistent carriage (Verwer et al., 2012). This suggests, that even without constituting the predominant clone in a hospital setting, ST78-MRSA-IV colonisation may be particularly likely to persist, a phenomenon which may have contributed the continued spread of this strain to H1 NICU patients over the eight-year study period.

A second CC88-MRSA clone, ST88-MRSA-IVa, was also identified in Irish hospitals during the present study. In Ireland, patients with specific HCA-MRSA risk factors generally undergo screening for MRSA (Irish Department of Health, 2013). These guidelines likely formed the basis upon which three ST88-MRSA-IVa isolates were recovered from three different patients during the study period. However, considering both the lack of epidemiological links between these isolates and more importantly, the high number of alleles by which they differed, it was concluded that this strain was introduced into Irish hospitals on three separate occasions (**Figure 1**). Moreover, the non-neonatal status of these patients further supported the likelihood of their having acquired this strain (generally considered CA) outside of a healthcare setting, prior to admission.

Extensive genotypic, conventional epidemiological and previously published data were all considered while determining the region(s) from which ST88-MRSA-IVa was likely imported into Ireland. Firstly, cgMLST indicated that two Irish ST88-MRSA-IVa isolates, P9, and P25, shared ancestral genotypes with isolates recovered in Egypt and France, respectively (**Figure 2**). Furthermore, sequence-based plasmid analysis revealed that all Irish and international ST88-MRSA-IVa isolates harboured the same *blaZ* and *cadX*-encoding plasmid, previously sequenced from a Tanzanian MRSA isolate. This suggested that all ST88-MRSA-IVa investigated may have originated in relatively close geographic proximity. Moreover, an Irish ST88-MRSA-IVa isolate harboured an additional plasmid, previously sequenced from a Ghanaian *S. aureus* isolate. Secondly, ST88 MRSA has become increasingly associated with Africa in recent years and France is known to have a large population of African ethnic origin (Schaumburg et al., 2014; https://www.insee.fr/en/statistiques/1283070). Finally, two of the three patients from whom ST88-MRSA-IVa was recovered, had African names, suggesting they may have had family connections to an African country. Considering these points, it was concluded that ST88-MRSA-IVa was likely imported into Ireland from Africa and/or a country with a large population of African ethnic origin.

While WGS and DNA microarray profiling were successfully utilised to achieve the aims of the present study, two significant limitations, which often impede WGS-based studies, were also identified. Firstly, in the absence of universally accepted intra-host strain diversity guidelines, putative transmission events were not identified exhaustively. Notably, however, isolates recovered

from twins (who generally share nursing care) on the same day, differed by eight alleles. Furthermore, isolates P10 and P11, which were recovered 3 days apart following a 30-month period in which the outbreak strain was not detected, differed by nine alleles. This suggests that an approximate intra-host strain diversity threshold of nine alleles may be applicable to the present study. Secondly, a lack of detailed epidemiological information limited the certainty with which conclusions could be drawn regarding the intricate dynamics of the H1 NICU outbreak, thus highlighting the importance of strong communicative links between healthcare facilities and research groups. However, despite the lack of detailed epidemiological information and the long period of time over which isolates were recovered, WGS provided robust and precise evidence of the occurrence of a protracted outbreak. Finally, this study highlighted the importance of considering all available epidemiological and genotypic information while selecting the whole-genome analysis approach best suited to the specific data set in question.

The present study revealed the HCW-facilitated spread of an Australian CA-MRSA strain, ST78-MRSA-IVa, in the NICU of an Irish maternity hospital over an eight-year period. Such findings indicate that further consideration of the role of HCWs in the transmission of MRSA in high-dependency units, such as NICUs, may be beneficial. This study also identified multiple introductions of an African CA-MRSA clone, ST88-MRSA-IVa, into Irish hospitals, suggesting that CA-MRSA risk factors should be considered during targeted patient screening. In a broader context, this study highlighted both the significance of travel in the spread of MRSA and the need for well-designed WGS-based studies that include in-depth epidemiological information in order to aid the establishment of data interpretation guidelines and thus, facilitate the real-time application of WGS in a clinical setting.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

and Dr. Helmut Hotzel (Institute of Bacterial Infections and Zoonoses, FLI, Jena, Germany) for providing the Australian, French and Egyptian/Tanzanian isolates, respectively. Finally, we thank Antje Ruppelt-Lorz, who in conjunction with SM, provided the German isolates.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01485/full#supplementary-material

REFERENCES

Australian Government Department of Immigration and Citizenship (2011). *Working Holiday Maker Visa Program Report*. Available online at: www.immi.gov.au/media/statistics/visitor.htm

Azarian, T., Cook, R. L., Johnson, J. A., Guzman, N., Mccarter, Y. S., Gomez, N., et al. (2016). Whole-genome sequencing for outbreak investigations of methicillin-resistant *Staphylococcus aureus* in the neonatal intensive care unit: time for routine practice? *Infect. Control Hosp. Epidemiol.* 36, 777–785. doi: 10.1017/ice.2015.73

Baig, S., Johannesen, T. B., Overballe-Petersen, S., Larsen, J., Larsen, A. R., and Stegger, M. (2018). Novel SCC *mec* type XIII (9A) identified in an ST152 methicillin-resistant *Staphylococcus aureus*. *Infect. Genet. Evol.* 61, 74–76. doi: 10.1016/j.meegid.2018.03.013

Bal, A. M., Coombs, G. W., Holden, M. T. G., Lindsay, J. A., Nimmo, G. R., Tattevin, P., et al. (2016). Genomic insights into the emergence and spread of international clones of healthcare- , community- and livestock-associated meticillin- resistant *Staphylococcus aureus*: blurring of the traditional definitions. *J. Glob. Antimicrob. Resist.* 6, 95–101. doi: 10.1016/j.jgar.2016.04.004

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Becker, K., van Alen, S., Idelevich, E. A., Schleimer, N., Seggewi,ß, J., Mellmann, A., et al. (2018). Plasmid-encoded transferable *mecB*-mediated methicillin resistance in *Staphylococcus aureus*. *Emerg. Infect. Dis.* 24, 242–248. doi: 10.3201/eid2402.171074

Brennan, G. I., Shore, A. C., Corcoran, S., Tecklenborg, S., Coleman, D. C., and O'Connella, B. (2012). Emergence of hospital- and community-associated Panton-Valentine leukocidin-positive methicillin-resistant *Staphylococcus aureus* genotype ST772-MRSA-V in Ireland and detailed investigation of an ST772-MRSA-V cluster in a neonatal intensive care unit. *J. Clin. Microbiol.* 50, 841–847. doi: 10.1128/JCM.06354-11

Breurec, S., Zriouil, S. B., Fall, C., Boisier, P., Brisse, S., Djibo, S., et al. (2011). Epidemiology of methicillin-resistant *Staphylococcus aureus* lineages in five major African towns: emergence and spread of atypical clones. *Clin. Microbiol. Infect.* 17, 160–165. doi: 10.1111/j.1469-0691.2010.03219.x

Coombs, G., Pearson, J., Robinson, O., and Christiansen, K. (2012). *Western Australian Methicillin-Resistant Staphylococcus Aureus (MRSA) and Vancomycin Resistant Enterococcus (VRE) Epidemiology and Typing Report*. Available online at: http://www.public.health.wa.gov.au/cproot/5576/2/access-annual-report-2012-13.pdf

Coombs, G. W., Daley, D. A., Pearson, J. C., Nimmo, G. R., Collignon, P. J., and Mclaws, M. (2014). Community-onset *Staphylococcus aureus* surveillance programme annual report, 2012. *Commun. Dis. Intell.* 38, 59–69.

Earls, M. R., Kinnevey, P. M., Brennan, G. I., Lazaris, A., Skally, M., O'Connell, B., et al. (2017). The recent emergence in hospitals of multidrug-resistant community-associated sequence type 1 and *spa* type t127 methicillin-resistant *Staphylococcus aureus* investigated by whole-genome sequencing: Implications for screening. *PLoS ONE* 12, 1–17. doi: 10.1371/journal.pone.0175542

European Committee on Antimicrobial Susceptibility Testing (2017). *Breakpoint tables for interpretation of MICs and zone diameters. Version 7.1*. Available online at: http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_7.1_Breakpoint_Tables.pdf

Feil, E. J., Cooper, J. E., Grundmann, H., Robinson, D. A., Enright, M. C., Berendt, T., et al. (2003). How clonal is *Staphylococcus aureus*? *J. Bacteriol.* 185, 3307–3316. doi: 10.1128/JB.185.11.3307-3316.2003

Frénay, H. M., Bunschoten, A. E., Schouls, L. M., van Leeuwen, W. J., Vandenbroucke-Grauls, C. M., Verhoef, J., et al. (1996). Molecular typing of methicillin-resistant *Staphylococcus aureus* on the basis of protein A gene polymorphism. *Eur. J. Clin. Microbiol. Infect. Dis.* 15, 60–64. doi: 10.1007/BF01586186

Geva, A., Wright, S. B., Baldini, L. M., Smallcomb, J. A., Safran, C., and Gray, J. E. (2011). Spread of methicillin-resistant *Staphylococcus aureus* in a large tertiary NICU: network analysis. *Pediatrics* 128, e1173–e1180. doi: 10.1542/peds.2010-2562

Irish Department of Health (2013). *Prevention and Control Methicillin-Resistant Staphylococcus aureus (MRSA) National Clinical Guideline* No.2.

Irish National Meticillin-Resistant *Staphylococcus aureus* Reference Laboratory (2016). *Annual Report 2016*.

Katayama, Y., Ito, T., and Hiramatsu, K. (2000). A new class of genetic element, staphylococcus cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 44, 1549–1555. doi: 10.1128/AAC.44.6.1549-1555.2000

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436

Kinnevey, P. M., Shore, A. C., Brennan, G. I., Sullivan, D. J., Ehricht, R., Monecke, S., et al. (2014). Extensive genetic diversity identified among sporadic methicillin-resistant *Staphylococcus aureus* isolates recovered in Irish hospitals between 2000 and 2012. *Antimicrob. Agents Chemother.* 58, 1907–1917. doi: 10.1128/AAC.02653-13

Leopold, S. R., Goering, R. V., Witten, A., Harmsen, D., and Mellmann, A. (2014). Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J. Clin. Microbiol.* 52, 2365–2370. doi: 10.1128/JCM.00262-14

Lindsay, J. A. (2010). Genomic variation and evolution of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 300, 98–103. doi: 10.1016/j.ijmm.2009.08.013

Malachowa, N., and Deleo, F. R. (2010). Mobile genetic elements of *Staphylococcus aureus*. *Cell. Mol. Life Sci.* 67, 3057–3071. doi: 10.1007/s00018-010-0389-4

McManus, B. A., Coleman, D. C., Deasy, E. C., Brennan, G. I., Connell, B. O., Monecke, S., et al. (2015). Comparative genotypes, staphylococcal cassette chromosome *mec* (SCC*mec*) genes and antimicrobial resistance amongst *Staphylococcus epidermidis* and *Staphylococcus haemolyticus* isolates from infections in humans and companion animals. *PLoS ONE* 10, 1–18. doi: 10.1371/journal.pone.0138079

Monecke, S., Coombs, G., Shore, A. C., Coleman, D. C., Akpaka, P., Borg, M., et al. (2011). A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. *PLoS ONE* 6:e0017936. doi: 10.1371/journal.pone.0017936

Monecke, S., Jatzwauk, L., Müller, E., Nitschke, H., Pfohl, K., Slickers, P., et al. (2016). Diversity of SCC*mec* elements in *Staphylococcus aureus* as observed in south-eastern Germany. *PLoS ONE* 11:e0162654. doi: 10.1371/journal.pone.0162654

Monecke, S., Skakni, L., Hasan, R., Ruppelt, A., Ghazal, S. S., Hakawi, A., et al. (2012). Characterisation of MRSA strains isolated from patients in a hospital in Riyadh, Kingdom of Saudi Arabia. *BMC Microbiol.* 12:146. doi: 10.1186/1471-2180-12-146

Monecke, S., Slickers, P., and Ehricht, R. (2008). Assignment of *Staphylococcus aureus* isolates to clonal complexes based on microarray analysis and pattern recognition. *FEMS Immunol. Med. Microbiol.* 53, 237–251. doi: 10.1111/j.1574-695X.2008.00426.x

Neely, A. N., and Maley, M. P. (2000). Survival of enterococci and staphylococci on hospital fabrics and plastics. *J. Clin. Microbiol.* 38, 724–726.

O'Brien, F. G., Coombs, G. W., Pearman, J. W., Gracey, M., Moss, F., Christiansen, K. J., et al. (2009). Population dynamics of methicillin-susceptible and -resistant

Staphylococcus aureus in remote communities. *J. Antimicrob. Chemother.* 64, 684–693. doi: 10.1093/jac/dkp285

Orlek, A., Stoesser, N., Anjum, M. F., Doumith, M., Ellington, M. J., Peto, T., et al. (2017). Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front. Microbiol.* 8:182. doi: 10.3389/fmicb.2017.00182

Ozaki, K., Takano, M., Higuchi, W., Takano, T., Yabe, S., Nitahara, Y., et al. (2009). Genotypes, intrafamilial transmission, and virulence potential of nasal methicillin-resistant *Staphylococcus aureus* from children in the community. *J. Infect. Chemother.* 15, 84–91. doi: 10.1007/s10156-009-0668-X

Parisi, A., Caruso, M., Normanno, G., Latorre, L., Sottili, R., Miccolupo, A., et al. (2016). Prevalence, antimicrobial susceptibility and molecular typing of methicillin-resistant *Staphylococcus aureus* (MRSA) in bulk tank milk from southern Italy. *Food Microbiol.* 58, 36–42. doi: 10.1016/j.fm.2016.03.004

Price, J. R., Didelot, X., Crook, D. W., Llewelyn, M. J., and Paul, J. (2013). Whole genome sequencing in the prevention and control of *Staphylococcus aureus* infection. *J. Hosp. Infect.* 83, 14–21. doi: 10.1016/j.jhin.2012.10.003

Price, J. R., Golubchik, T., Cole, K., Wilson, D. J., Crook, D. W., Thwaites, G. E., et al. (2014). Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of *Staphylococcus aureus* in an intensive care unit. *Clin. Infect. Dis.* 58, 609–618. doi: 10.1093/cid/cit807

Qiao, Y., Xue, N., Qiang, C., Zhao, R., Song, W., Zheng, Y., et al. (2014). Clinical and molecular characteristics of invasive community-acquired *Staphylococcus aureus* infections in Chinese children. *Bmc Infect. Dis.* 14:582. doi: 10.1186/s12879-014-0582-4

Raji, M. A., Garaween, G., Ehricht, R., Monecke, S., Shibl, A. M., and Senok, A. (2016). Genetic characterization of *Staphylococcus aureus* isolated from retail meat in Riyadh, Saudi Arabia. *Front. Microbiol.* 7:911. doi: 10.3389/fmicb.2016.00911

Robinson, D. A., and Enright, M. C. (2004). Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 10, 92–97. doi: 10.1111/j.1469-0691.2004.00768.x

Roisin, S., Gaudin, C., De Mendonça, R., Bellon, J., Van Vaerenbergh, K., De Bruyne, K., et al. (2016). Pan-genome multilocus sequence typing and outbreak-specific reference-based single nucleotide polymorphism analysis to resolve two concurrent *Staphylococcus aureus* outbreaks in neonatal services. *Clin. Microbiol. Infect.* 22, 520–526. doi: 10.1016/j.cmi.2016.01.024

RuŽičková, V., Pantuček, R., Petráš, P., Machová, I., Kostýlková, K., and Doškar, J. (2012). Major clonal lineages in impetigo *Staphylococcus aureus* strains isolated in Czech and Slovak maternity hospitals. *Int. J. Med. Microbiol.* 302, 237–241. doi: 10.1016/j.ijmm.2012.04.001

Sabat, A. J., Hermelijn, S. M., Akkerboom, V., Juliana, A., Degener, J. E., Grundmann, H., et al. (2017). Complete-genome sequencing elucidates outbreak dynamics of CA-MRSA USA300 (ST8-*spa* t008) in an academic hospital of Paramaribo, Republic of Suriname. *Sci. Rep.* 7, 1–12. doi: 10.1038/srep41050

Schaumburg, F., Alabi, A. S., Peters, G., and Becker, K. (2014). New epidemiology of *Staphylococcus aureus* infection in Africa. *Clin. Microbiol. Infect.* 20, 589–596. doi: 10.1111/1469-0691.12690

Schaumburg, F., Ngoa, U. A., Kösters, K., Köck, R., Adegnika, A. A., Kremsner, P. G., et al. (2011). Virulence factors and genotypes of *Staphylococcus aureus* from infection and carriage in Gabon. *Clin. Microbiol. Infect.* 17, 1507–1513. doi: 10.1111/j.1469-0691.2011.03534.x

Schürch, A. C., Arredondo-Alonso, S., Willems, R. J. L., and Goering, R. V. (2018). Whole genome sequencing (WGS) options for bacterial strain typing and epidemiological analysis based on single nucleotide polymorphism (SNP) vs gene-by-gene-based approaches. *Clin. Microbiol. Infect.* 24, 350–354. doi: 10.1016/j.cmi.2017.12.016

Shore, A. C., and Coleman, D. C. (2013). Staphylococcal cassette chromosome *mec*: recent advances and new insights. *Int. J. Med. Microbiol.* 303, 350–359. doi: 10.1016/j.ijmm.2013.02.002

Shore, A. C., Deasy, E. C., Slickers, P., Brennan, G., O'Connell, B., Monecke, S., et al. (2011). Detection of staphylococcal cassette chromosome *mec* type XI carrying highly divergent *mecA*, *mecI*, *mecR1*, *blaZ*, and *ccr* genes in human clinical isolates of clonal complex 130 methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 55, 3765–3773. doi: 10.1128/AAC.00187-11

Verwer, P. E. B., Robinson, J. O., Coombs, G. W., Wijesuriya, T., Murray, R. J., Verbrugh, H. A., et al. (2012). Prevalence of nasal methicillin-resistant *Staphylococcus aureus* colonization in healthcare workers in a Western Australian acute care hospital. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 1067–1072. doi: 10.1007/s10096-011-1408-6

Vindel, A., Trincado, P., Cuevas, O., Ballesteros, C., Bouza, E., and Cercenado, E. (2014). Molecular epidemiology of community-associated methicillin-resistant *Staphylococcus aureus* in Spain: 2004-12. *J. Antimicrob. Chemother.* 69, 2913–2919. doi: 10.1093/jac/dku232

Waldron, D. E., and Lindsay, J. A. (2006). Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between *S. aureus* isolates of different lineages. *J. Bacteriol.* 188, 5578–5585. doi: 10.1128/JB.00418-06

Wesley Long, S., Beres, S. B., Olsen, R. J., and Musser, J. M. (2014). Absence of patient-to-patient intrahospital transmission of *Staphylococcus aureus* as determined by whole-genome sequencing. *MBio* 5, 1–10. doi: 10.1128/mBio.01692-14

Wu, Z., Li, F., Liu, D., Xue, H., and Zhao, X. (2015). Novel type XII staphylococcal cassette chromosome *mec* harboring a new cassette chromosome recombinase, CcrC2. *Antimicrob. Agents Chemother.* 59, 7597–7601. doi: 10.1128/AAC.01692-15

Young, B. C., Golubchik, T., Batty, E. M., Fung, R., Larner-svensson, H., Rimmer, A. J., et al. (2012). Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4550–4555. doi: 10.1073/pnas.1113219109

Zhang, W., Shen, X., Zhang, H., Wang, C., Deng, Q., Liu, L., et al. (2009). Molecular epidemiological analysis of methicillin-resistant *Staphylococcus aureus* isolates from Chinese pediatric patients. *Eur. J. Clin. Microbiol. Infect. Dis.* 28, 861–864. doi: 10.1007/s10096-009-0706-8

Check for updates

A Sporadic Four-Year Hospital Outbreak of a ST97-IVa MRSA With Half of the Patients First Identified in the Community

Ingrid M. Rubin[1], Thomas A. Hansen[1], Anne Mette Klingenberg[1], Andreas M. Petersen[1,2,3], Peder Worning[1], Henrik Westh[1,3] and Mette D. Bartels[1]*

[1] Department of Clinical Microbiology, Hvidovre Hospital, Hvidovre, Denmark, [2] Department of Gastroenterology, Hvidovre Hospital, Hvidovre, Denmark, [3] Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

This study describes a sporadically occurring 4-year outbreak of methicillin-resistant *Staphylococcus aureus* (MRSA) originating from a surgical ward. Whole-genome sequencing (WGS) identified the outbreak clone as *spa* type t267, sequence type ST97, and SCC*mec* IVa. Prompted by the finding of four patients within 6 months in the same ward with this unusual MRSA type, an outbreak was suspected. Subsequent MRSA screening in the ward in February 2017 identified three-additional patients and two health care workers (HCWs) with t267/ST97-IVa. WGS linked these 9 isolates to 16 previous isolates in our WGS database and the outbreak thus included 23 patients and two HCWs. Twenty-one patients had a connection to the surgery ward during the period 2013–2017, but half of them had MRSA diagnosed in the community long after discharge. The community debut of several patients MRSA infections weeks to months after hospital discharge made the identification of a hospital source difficult and it was the SNP relatedness of the isolates that led us to identify the common denominator of hospitalization. An index patient was not identified, but our hypothesis is that HCWs with unrecognized long-term MRSA colonization could have caused sporadic nosocomial transmission due to intermittent breaches in infection prevention and control practice.

Keywords: WGS, outbreak, CO-MRSA, ST97, HCWs

INTRODUCTION

Methicillin-resistant *Staphylococcus aureus* (MRSA) has been a global medical challenge since its emergence in 1961, 2 years after methicillin was introduced to treat penicillin-resistant *S. aureus* (Jevons, 1961). In human medicine, the focus has traditionally been on the hospital-acquired clones (HA-MRSA), but new clones have emerged both in livestock and in community settings (DeLeo et al., 2010; Gonçalves da Silva et al., 2017). Outbreaks in hospitals and nursing homes are today caused by both HA-MRSA and community-associated MRSA (CA-MRSA) clones (DeLeo et al., 2010; Di Ruscio et al., 2017; Henderson and Nimmo, 2017) and it has been suggested no longer to regard HA-MRSA and CA-MRSA as separate entities (Zarfel et al., 2013). Hospital admission of unknown MRSA carriers, lack of MRSA admittance screening and spread of MRSA among nursing home residents could all be part of the explanation of this mélange of HA-MRSA and CA-MRSA in hospitals (Gonzalez et al., 2006). General contact procedures, an important part of infection

control, will keep patients safe from transmission of MRSA from health care workers (HCWs) (Harris A.D. et al., 2013, Harris et al., 2017). By typing MRSA isolates and drawing interferences on transmission based on genetic relatedness, transmission pathways can be tracked (Harris S.R. et al., 2013). Today the most precise typing of bacteria is by whole-genome sequencing (WGS). WGS has proven an excellent tool, with good inter-laboratory reproducibility in hospital outbreaks (Bartels et al., 2013; Leopold et al., 2014; SenGupta et al., 2014).

Here, we describe a prolonged outbreak of MRSA initiated in a hospital ward that was confirmed by WGS. Tracing of the clone led us to 23 patients and two HCWs, who in most cases had a common denominator in one of the hospitals surgery wards.

MATERIALS AND METHODS

Setting
This retrospective study analyzed an outbreak spanning the period between June 2013 and February 2017 that occurred at Hvidovre Hospital, Copenhagen, Denmark. The MRSA isolates studied were routinely found in clinical samples or after the outbreak was discovered in February 2017 through MRSA screening of patients and staff.

Data Set
All MRSA isolates are investigated by WGS as part of the routine at the Department of Clinical Microbiology, Hvidovre Hospital. In order to define the outbreak, we investigated the relatedness to other t267/ST97/SCC*mec* IVa isolates from the period 2013–2017 in the Capital Region of Denmark and these isolates were included in the SNP analysis.

Whole-Genome Sequencing and Analysis
Each MRSA isolate was initially confirmed with an in-house multiplex real-time polymerase chain reaction (PCR) that detects the presence of *nuc*, *fem*A, *mec*A, and *mec*C. Since January 2013 all MRSA isolates have been WGS on a MiSeq (Illumina, United States). DNA extraction were performed on all MRSA isolates and libraries were made with 2 × 150 bp paired-end Nextera XT DNA sample preparation kit (Illumina, United States) and sequenced on a MiSeq (Illumina, United States). The reads were mapped to a USA300 reference sequence (US300_TCH1516) using stampy (Lunter and Goodson, 2011) with an expected substitution rate of 0.01 (Didelot et al., 2012) for single nucleotide variants detection. Variants were called using SAMtools v0.1.12 (Li, 2011) mpileup command with options -M0 -Q30 -q30 -o40 -e20 -h100 -m2 -D -S. The genome was assembled using Velvet v1.0.11 (Zerbino and Birney, 2008) or Spades (Bankevich et al., 2012). Phylogeny was inferred by neighbor-joining analysis.

Ethical Considerations
Permission to link the sequencing of MRSA from routine clinical samples to patient data without individual patient consent was obtained from the Danish Data Protection Agency

(no. AHH-2017-095, I-Suite nr. 06029). Permission to look up the patients admission data without individual consent was obtained from the Hospital Board (no. WZ17038300-2018-14).

RESULTS

Documentation of the Outbreak
In January 2017, an MRSA infected patient was found in one of the hospital's surgery wards. From the sample of abdominal pus two distinct MRSA types were found based on antimicrobial susceptibility patterns. WGS of the two isolates identified a t267/ST97-IVa and a t002/ST5-IVg. Looking 6 months back in our MRSA WGS database, we identified three other patients admitted to the same ward that had tested positive for MRSA t267/ST97. Searching our WGS database back to 2013 we found this to be an unusual MRSA type in our region with just 34 isolates from 2013 to January 2017 (0.7%). A neighbor-joining tree was then constructed with all 34 isolates showing 20 isolates with a close connection of ≤50 SNPs. These isolates also shared the same SCC*mec* type, namely IVa. Thus, the suspicion of an MRSA outbreak was confirmed. This led to a several week-long screening of patients, HCWs and other staff members in the ward resulting in the finding of three-additional patients and two HCWs with t267/ST97-IVa. Furthermore, screening at the ward revealed that two HCWs carried other MRSA types (t045/ST5-IVc/e, t002/ST5-IVg) and one patient with known contact to pigs carried the livestock-associated t034/ST398-V.

In our final outbreak cluster (**Figure 1**) we ended up with 25 persons. Eighteen had been admitted to the surgical ward, two were HCWs, two were family members of MRSA positive patients, and one had shared a room at another ward in another hospital with an MRSA positive patient, who had previously been admitted to the surgery ward. Finally, two patients had no relation either to another patient or to the ward to the best of our knowledge. We have data on hospitalization at the surgery ward for most patients but the surgery ward has many sub-departments and we lack data on room allocations as well as the exact department admitted to. Our data show that 13 patients had an overlap in admission period with at least one other patient (**Supplementary Figure S1**).

Nine patients had been diagnosed with MRSA t267/ST97-IVa at their General Practitioner and three more had been diagnosed at the Emergency Room, resulting in 12 cases first identified in the community, here defined as community-onset MRSA (CO-MRSA). Eight patients had their MRSA diagnosed at hospital wards HA-MRSA. Five persons had been diagnosed by screening of the hospital ward (including the two HCWs) (**Table 1**).

Due to the finding of other MRSA types among our HCWs, we also studied the phylogenetic tree of all the t045/ST5-IVc/e and t002/ST5-IVg isolates in our WGS database. There were only three t045/ST5 isolates and the two others had another SCC*mec* type (**Supplementary Figure S2**). Our database

FIGURE 1 | Neighbor-joining tree for the 25 isolates in our outbreak cluster with t267/ST97 SCC*mec* IVa. Scale bar indicates the SNP distance. Green: HCWs, Red: patients admitted in 2016 and 2017, Black: patients admitted 2013–2015, Blue: patients with no identified relation to the surgical ward.
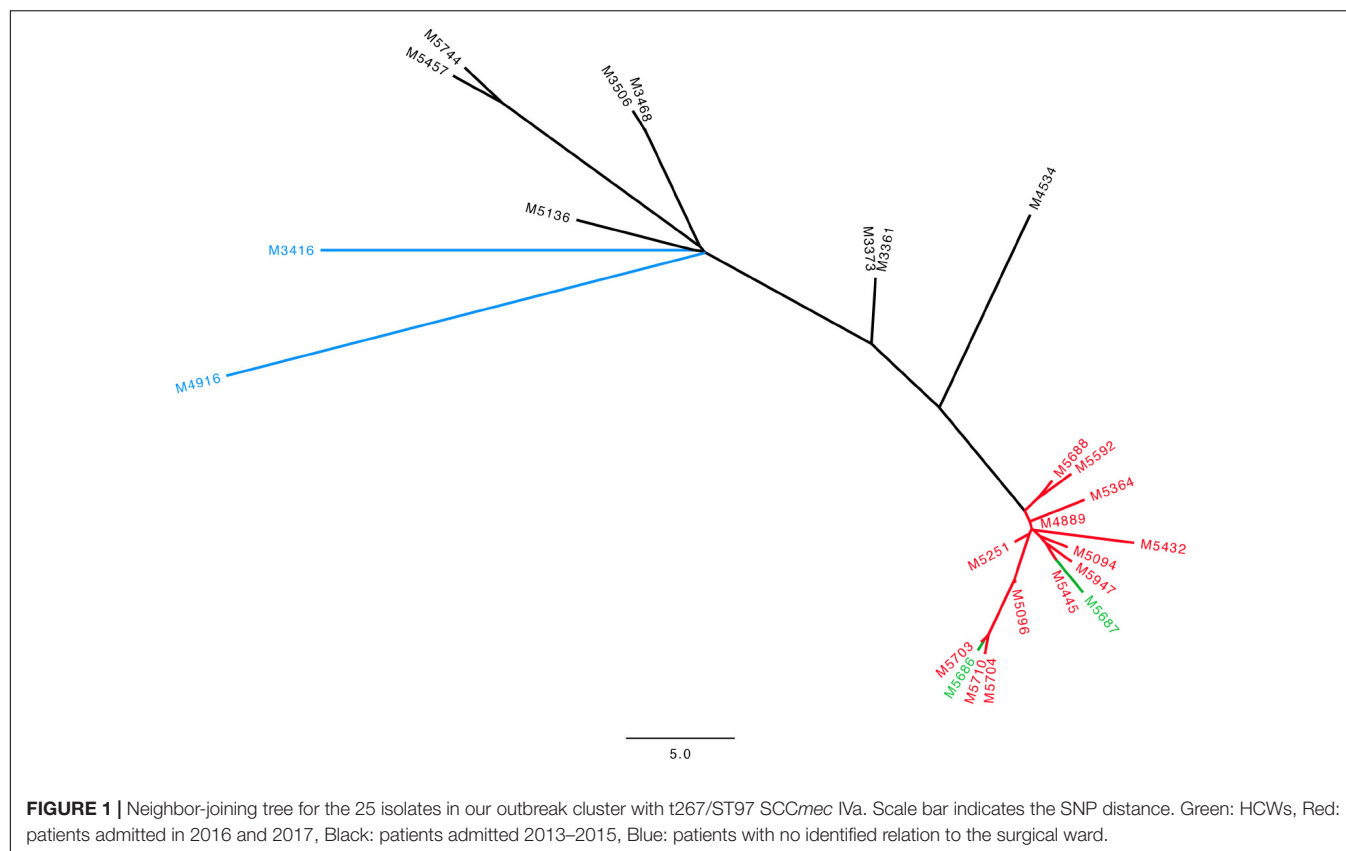
contained 266 t002/ST5 isolates of which 40 harbored SCC*mec* IVg (**Supplementary Figure S3**).

SNP Analysis and Characterization of the Outbreak Clone

The 25 t267/ST97/IVa isolates showed a low diversity with a maximum of 50 SNP differences over the 4-year period. Within our outbreak cluster, a further sub-cluster could be distinguished with 15 isolates with a maximum of 11 SNPs. This sub-cluster consisted of 13 patients admitted in 2016 and 2017 as well as the two HCWs.

All outbreak isolates belonged to *spa* type t267 and ST97 except for two isolates with an unnamed ST that was a single locus variant (SLV) of ST97. All isolates had SCC*mec* IVa and no isolates had Panton–Valentine leukocidin (PVL) or the arginine catabolic mobile element (ACME). The 14 non-outbreak related isolates of t267/ST97 had SCC*mec* IVc/e (Gonzalez et al., 2006), SCC*mec* IVa (Henderson and Nimmo, 2017), and SCC*mec* V (Gonçalves da Silva et al., 2017). The outbreak isolates were resistant to methicillin and susceptible to erythromycin, clindamycin, gentamicin, fusidic acid, linezolid, mupirocin, trimethoprim/sulfamethoxazole, and rifampicin.

Infection Control Measures

An outbreak group with representatives from the surgical ward, the Department of Cleaning, the Department of Clinical Microbiology, and the Infection Control Organization was

established. In order to exclude shortcomings of general infection control precautions, various initiatives were launched. The ward had daily visits by the infection control and prevention nurse, where behavior was observed and adherence to procedures monitored, and on this basis an increased focus was placed on the use of protective equipment such as gloves and plastic aprons, which among other things led to an increased availability of protective equipment in the department. Furthermore, focus was on infection control precautions for both HCWs and patients and in particular on how patients could be motivated for better hand hygiene.

The ward was cleaned by standard hospital cleaning, followed by manual disinfection with bleach. The HCWs with the outbreak clone were successfully decolonized.

DISCUSSION

It is a global trend that patients are hospitalized for increasingly shorter time periods. Therefore, hospital acquisition of an MRSA might not be suspected or identified due to clinical onset long after discharge. Healthcare-associated MRSA outbreaks are rare in Denmark (Andersen and Knudse, 2016) and have been predominantly associated with outbreaks in Neonatal Wards (Ramsing et al., 2013; Bartels et al., 2015; Franck et al., 2017). The extent to which the community serves as a reservoir to the spread of MRSA into hospitals is largely unknown, but globally several reports on CA-MRSA in hospitals have emerged (DeLeo et al.,

TABLE 1 | Demographics.

Patients	Time of MRSA positive	Year admitted to the surgery ward	Sites of colonization	Hospital-acquired vs. community-onset	Comment
M5592	January 2017	2017	Abdominal pus	HA	
M5703	February 2017	2017	Screening ward	HA	Shared room with M5704
M5704	February 2017	2017	Screening ward	HA	Shared room with M5703
M5688	February 2017	2017	Screening ward	HA	
M5710	February 2017	2016 and 2017	Wound	CO	
M5744	March 2017	2016	Tracheal secretion	CO	
M5445	November 2016	2013 and 2016	Cicatrice	HA	
M5432	November 2016	2016	Urine	CO	
M5251	September 2016	2016	Catheter	HA	
M5136	July 2016	2016	Urine	HA	
M5096	July 2016	2016	Blood	HA	
M5094	June 2016	2016	Absces	CO	
M4889	March 2016	2016	Cicatrice	CO	
M4534	September 2015	2015 and 2016	Cicatrice	CO	
M5947	June 2017	2015 and 2016	Wound	CO	
M5364	October 2016	2015	Cicatrice	CO	
M3506	March 2014		Screening	HA	Screening due to MRSA positive household contact (M3468)
M3468	February 2014	2014	Wound	CO	Household contact with M3506
M3361	December 2013	2013	Nose	HA	
M5687	February 2017		Screening ward		HCW
M5686	February 2017		Screening ward		HCW
M5457	November 2016		Screening	CO	Screening due to MRSA positive household contact (M4744)
M3373	December 2013		Screening	HA	Shared room at another hospital with ++ M3361 in 2013
M4916	April 2016		Screening GP	CO	Unknown connection
M3416	January 2014		Screening GP	CO	Unknown connection

2010; Thurlow et al., 2012). With repeated introduction of CA-MRSA (Cho and Chung, 2017; Coll et al., 2017), a better action plan is needed to tackle and curb community-associated carriage (Bartels et al., 2010). If the level of MRSA carriage increases in the general population then it will also increase in patients and HCWs. In this situation the sporadic spread of MRSA between HCWs and patients with unknown carrier-state, that we find in low MRSA prevalence countries might contribute to more intermittent transmission despite general infection control procedures.

Here, we report a 4-year long outbreak of MRSA type t267/ST97 SCC*mec* IVa, that was discovered as the MRSA was found in four patients in the same ward within 6 months. This, for us, rare MRSA has sporadically been found around the world and has been described both as a CA-MRSA (Monecke et al., 2011) and as a LA-MRSA in pigs and associated with bovine mastitis (Menegotto et al., 2012; Pantosti, 2012; Feltrin et al., 2015). We have routinely WGS all MRSA isolates since January 2013 and a phylogenetic tree clustered 25 isolates together and patient records revealed that the common denominator was the surgery ward. The outbreak isolates differed up to 50 SNPs, with a subcluster of 13 patients from 2016 or 2017 and the two HCWs whose isolates differed by no more than 11 SNPs. This gives an evolution of the core genome of about 5–6 SNPS per year in our

sub-cluster comparable to the 6–9 SNPs per year described in one study (Holden et al., 2013) and more than the 3–4 SNPs per year in other studies (Harris et al., 2010; Senn et al., 2016). The remaining 14 isolates with the same MRSA type had no link to the ward in question and the SNP distance to the outbreak isolates was between 50 and 249 SNPs. Furthermore, most of these isolates had a different SCC*mec*, which also indicates another clone.

Since the outbreak was ongoing for such a prolonged period at one of our busiest surgical wards, the actual number of people infected or colonized with the MRSA clonal isolate is probably higher than the rather modest number we report. Only individuals in the community who had a clinical infection would have been identified through a sample so MRSA carriers can very well have eluded the system. There could be various explanations to the increase in the number of positive MRSA patients in 2016–2017 in this study. Of course, we found quite a few through the screening in 2017 that might not have been found if the screening had not been performed. Another aspect could be that there were unknown MRSA positive HCWs that had quit the ward before the screening in 2017, but contributed to the increase in cases in 2016 and 2017. Nevertheless, the finding of relatively few patients over a 4-year period indicates there is relatively little spread between HCWs and patients. Previous studies have concluded that

nosocomial outbreaks caused by HCWs represent rare events, and therefore screening of personnel should not be performed regularly (Danzmann et al., 2013). Another study proposes three possible scenarios on the role of HCWs: being vectors of transmission, persistent reservoirs, or innocent by-standers, and concludes by suggesting aggressive screening and eradication policies in outbreak investigations (Albrich and Harbarth, 2008). Another aspect of the role of HCWs, and one that is enhanced by our findings, is that due to good infection control measures, HCWs with unrecognized long-term colonization cause only sporadic transmission. This is further supported by the finding of two other MRSA types in the HCWs during the ward screening, with no documented outbreaks caused by them. In support of the conclusion that there had been bacterial transmission from HCWs to patients is the fact that after the HCWs were declared free of MRSA, no more outbreak isolates have been seen at the ward as of May 2018. Of course, transmission could also have occurred between patients, since 13 patients had an overlap in admission time with at least one other patient. However, we lack data on room and sub-department allocation to support this. Due to good infection control measurements transmission only occurred rarely as the outbreak was going on unnoticed.

To automatically detect potential outbreaks the combination of detailed epidemiological data together with WGS is crucial (Roer et al., 2017). In the future an electronic system linking a patients isolate to the last seen most related isolate, will enable earlier detection of outbreaks (Mellmann et al., 2016).

CONCLUSION

Whole-genome sequencing can enhance the detection of prolonged hospital outbreaks of MRSA. Furthermore, we bring to light the fact that increasingly shorter hospital stays delays outbreak detection if continuous analysis of WGS is not performed. In this study, we identified an MRSA outbreak in

February 2017 and with the use of WGS we could trace the outbreak 4 years back in time.

We hypothesize that HCWs with an unknown MRSA carrier-state might cause sporadic transmission and sustain an unknown outbreak over many years. However, we believe that HCWs with a known MRSA carrier-state usually do not cause transmission due to their personally increased awareness of the importance of infection control standard precautions.

AUTHOR CONTRIBUTIONS

TH and PW performed the bioinformatics analyses. All authors contributed to the writing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01494/full#supplementary-material

FIGURE S1 | Timeline of hospitalization dates at the surgery ward and date of first positive MRSA sample. Thirteen patients had an overlap with at least one other patient. M5136 had an admission in June 2016, but the exact dates were not available. Sample date denotes the date when MRSA was first discovered in the patient.

FIGURE S2 | Neighbor-joining tree for the three isolates with t045/ST5-IVc/e. Scale bar indicates the SNP distance. M5696 is the HCW. The difference between the isolates was more than 200 SNPs.

FIGURE S3 | NJ-tree of the t002/ST5-IVg isolates. Scale bar indicates the SNP distance. Of the 40 isolates we choose to portray the 25 isolates with less than 100 SNPs difference. The isolates connected with the HCW at the surgical ward are highlighted and the isolate of the HCW in question is M5667. M6160 is a household contact. RH410 is the patient she nursed at the ward, who in turn was also positive for t267, and part of our outbreak cluster. M5746 is an unknown connection, but with no known link to the surgery ward. Between these four isolates there are four SNPs at the most.

REFERENCES

Albrich, W. C., and Harbarth, S. (2008). Health-care workers: source, vector, or victim of MRSA? *Lancet Infect. Dis.* 8, 289–301. doi: 10.1016/S1473-3099(08)70097-5

Andersen, P. H., and Knudse, L. K. (2016). *MRSA*. Available at: https://www.ssi.dk/Aktuelt/Nyhedsbreve/EPI-NYT/2017/Uge%2023%20-%202017.aspx.2017

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bartels, M. D., Boye, K., Oliveira, D. C., Worning, P., Goering, R., and Westh, H. (2013). Associations between dru types and SCCmec cassettes. *PLoS One* 8:e61860. doi: 10.1371/journal.pone.0061860

Bartels, M. D., Kristoffersen, K., Boye, K., and Westh, H. (2010). Rise and subsequent decline of community-associated methicillin resistant *Staphylococcus aureus* ST30-IVc in Copenhagen, Denmark through an effective search and destroy policy. *Clin. Microbiol. Infect.* 16, 78–83. doi: 10.1111/j.1469-0691.2009.02829.x

Bartels, M. D., Larner-Svensson, H., Meiniche, H., Kristoffersen, K., Schønning, K., Nielsen, J. B., et al. (2015). Monitoring meticillin resistant *Staphylococcus aureus* and its spread in Copenhagen, Denmark, 2013, through routine whole

genome sequencing. *Euro Surveill.* 20:21112. doi: 10.2807/1560-7917.ES2015.20.17.21112

Cho, S. Y., and Chung, D. R. (2017). Infection prevention strategy in hospitals in the era of community-associated methicillin-resistant *Staphylococcus aureus* in the Asia-Pacific region: a review. *Clin. Infect. Dis.* 64, S82–S90. doi: 10.1093/cid/cix133

Coll, F., Harrison, E. M., Toleman, M. S., Reuter, S., Raven, K. E., Blane, B., et al. (2017). Longitudinal genomic surveillance of MRSA in the UK reveals transmission patterns in hospitals and the community. *Sci. Transl. Med.* 9:eaak9745.

Danzmann, L., Gastmeier, P., Schwab, F., and Vonberg, R.-P. (2013). Health care workers causing large nosocomial outbreaks: a systematic review. *BMC Infect. Dis.* 13:98. doi: 10.1186/1471-2334-13-98

DeLeo, F. R., Otto, M., Kreiswirth, B. N. B., and Chambers, H. F. H. (2010). Community-associated meticillin-resistant *Staphylococcus aureus*. *Lancet* 375, 1557–1568. doi: 10.1016/S0140-6736(09)61999-1

Di Ruscio, F., Bjørnholt, J. V., Larssen, K. W., Leegaard, T. M., Moen, A. E. F., and de Blasio, B. F. (2017). Epidemiology and spa-type diversity of meticillin-resistant *Staphylococcus aureus* in the community and healthcare settings in Norway. *J. Hosp. Infect.* doi: 10.1016/j.jhin.2017.12.019 [Epub ahead of print].

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., and Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* 13, 601–612. doi: 10.1038/nrg3226

Feltrin, F., Alba, P., Kraushaar, B., Ianzano, A., Argudín, M. A., Di Matteo, P., et al. (2015). A livestock-associated, multidrug-resistant, methicillin-resistant *Staphylococcus aureus* clonal complex 97 lineage spreading in dairy cattle and pigs in Italy. *Appl. Environ. Microbiol.* 82, 816–821. doi: 10.1128/AEM. 02854-15

Franck, K. T., Gumpert, H., Olesen, B., Larsen, A. R., Petersen, A., Bangsborg, J., et al. (2017). *Staphylococcal aureus* Enterotoxin C and enterotoxin-like L associated with post-partum mastitis. *Front. Microbiol.* 8:173. doi: 10.3389/fmicb.2017.00173

Gonçalves da Silva, A., Baines, S. L., Carter, G. P., Heffernan, H., French, N. P., Ren, X., et al. (2017). A phylogenomic framework for assessing the global emergence and evolution of clonal complex 398 methicillin-resistant *Staphylococcus aureus*. *Microb. Genomics* 3:e000105. doi: 10.1099/mgen.0. 000105

Gonzalez, B. E., Rueda, A. M., Shelburne, S. A. III, Musher, D. M., Hamill, R. J., and Hulten, K. G. (2006). community-associated strains of methicillin-resistant *Staphylococccus aureus* as the cause of healthcare-associated infection. *Infect. Control Hosp. Epidemiol.* 27, 1051–1056. doi: 10.1086/507923

Harris, A. D., Morgan, D. J., Pineles, L., Perencevich, E. N., and Barnes, S. L. (2017). Deconstructing the relative benefits of a universal glove and gown intervention on MRSA acquisition. *J. Hosp. Infect.* 96, 49–53. doi: 10.1016/j.jhin.2017.03.011

Harris, A. D., Pineles, L., Belton, B., Johnson, J. K., Shardell, M., Loeb, M., et al. (2013). Universal glove and gown use and acquisition of antibiotic-resistant bacteria in the ICU: a randomized trial. *JAMA* 310, 1571–1580. doi: 10.1001/jama.2013.277815

Harris, S. R., Cartwright, E. J. P., Török, M. E., Holden, M. T. G., Brown, N. M., Ogilvy-Stuart, A. L., et al. (2013). Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect. Dis.* 13, 130–136. doi: 10.1016/S1473-3099(12)70268-2

Harris, S. R., Feil, E. J., Holden, M. T. G., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science. 1182395

Henderson, A., and Nimmo, G. R. (2017). Control of healthcare- and community-associated MRSA: recent progress and persisting challenges. *Br. Med. Bull.* 125, 1–17. doi: 10.1093/bmb/ldx046

Holden, M. T. G., Hsu, L.-Y., Kurt, K., Weinert, L. A., Mather, A. E., Harris, S. R., et al. (2013). A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res.* 23, 653–664. doi: 10.1101/gr.147710.112

Jevons, M. P. (1961). "Celbenin" -resistant staphylococci. *Br. Med. J.* 1, 124–125. doi: 10.1136/bmj.1.5219.124-a

Leopold, S. R., Goering, R. V., Witten, A., Harmsen, D., and Mellmann, A. (2014). Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J. Clin. Microbiol.* 52, 2365–2370. doi: 10.1128/JCM.00262-14

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Lunter, G., and Goodson, M. (2011). A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* 21, 936–939. doi: 10.1101/gr. 111120.110

Mellmann, A., Bletz, S., Böking, T., Kipp, F., Becker, K., Schultes, A., et al. (2016). Real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. *J. Clin. Microbiol.* 54, 2874–2881. doi: 10.1128/JCM.00790-16

Menegotto, F., González-Cabrero, S., Lorenzo, B., Cubero, Á., Cuervo, W., Gutiérrez, M. P., et al. (2012). Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in a Spanish hospital over a 4-year period: clonal replacement, decreased antimicrobial resistance, and identification of community-acquired and livestock-associated clones. *Diagn. Microbiol. Infect. Dis.* 74, 332–337. doi: 10.1016/j.diagmicrobio.2012.08.001

Monecke, S., Coombs, G., Shore, A. C., Coleman, D. C., Akpaka, P., Borg, M., et al. (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

Pantosti, A. (2012). Methicillin-resistant *Staphylococcus aureus* associated with animals and its relevance to human health. *Front. Microbiol.* 3:127. doi: 10.3389/fmicb.2012.00127

Ramsing, B. G. U., Arpi, M., Andersen, E. A., Knabe, N., Mogensen, D., Buhl, D., et al. (2013). First outbreak with mrsa in a danish neonatal intensive care unit: risk factors and control procedures. *PLoS One* 8:e66904. doi: 10.1371/journal. pone.0066904

Roer, L., Hansen, F., Frølund Thomsen, M. C., Knudsen, J. D., Hansen, D. S., Wang, M., et al. (2017). WGS-based surveillance of third-generation cephalosporin-resistant *Escherichia coli* from bloodstream infections in Denmark. *J. Antimicrob. Chemother.* 72, 1922–1929. doi: 10.1093/jac/dkx092

SenGupta, D. J., Cummings, L. A., Hoogestraat, D. R., Butler-Wu, S. M., Shendure, J., Cookson, B. T., et al. (2014). Whole-genome sequencing for high-resolution investigation of methicillin-resistant *Staphylococcus aureus* epidemiology and genome plasticity. *J. Clin. Microbiol.* 52, 2787–2796. doi: 10.1128/JCM.00759-14

Senn, L., Clerc, O., Zanetti, G., Basset, P., Prod'hom, G., Gordon, N. C., et al. (2016). The stealthy superbug: the role of asymptomatic enteric carriage in maintaining a long-term hospital outbreak of ST228 methicillin-resistant *Staphylococcus aureus*. *mBio* 7:e02039-15. doi: 10.1128/mBio.02039-15

Thurlow, L. R., Joshi, G. S., and Richardson, A. R. (2012). Virulence strategies of the dominant USA300 lineage of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA). *FEMS Immunol. Med. Microbiol.* 65, 5–22. doi: 10.1111/j.1574-695X.2012.00937.x

Zarfel, G., Luxner, J., Folli, B., Leitner, E., Feierl, G., Kittinger, C., et al. (2013). Genomic insights into the emergence and spread of international clones of healthcare-, community- and livestock-associated meticillin-resistant *Staphylococcus aureus*: blurring of the traditional definitions. *J. Clin. Microbiol.* 7, 1–10.

Zerbino, D. R., and Birney, E. (2008). Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Global Scale Dissemination of ST93: A Divergent *Staphylococcus aureus* Epidemic Lineage That Has Recently Emerged From Remote Northern Australia

Sebastiaan J. van Hal[1], Eike J. Steinig[2,3], Patiyan Andersson[2], Matthew T. G. Holden[4,5], Simon R. Harris[5], Graeme R. Nimmo[6], Deborah A. Williamson[7], Helen Heffernan[8], S. R. Ritchie[9], Angela M. Kearns[10], Matthew J. Ellington[11], Elizabeth Dickson[12], Herminia de Lencastre[13,14], Geoffrey W. Coombs[15,16], Stephen D. Bentley[5], Julian Parkhill[5], Deborah C. Holt[2], Phillip M. Giffard[2] and Steven Y. C. Tong[2,17]*

[1] Department of Microbiology and Infectious Diseases, Royal Prince Alfred Hospital, Sydney, NSW, Australia, [2] Global and Tropical Health Division, Menzies School of Health Research, Darwin, NT, Australia, [3] Australian Institute of Tropical Health and Medicine, Townsville, QLD, Australia, [4] School of Medicine, University of St. Andrews, Fife, United Kingdom, [5] Pathogen Genomics, Wellcome Trust Sanger Institute, Cambridge, United Kingdom, [6] Pathology Queensland Central Laboratory and Griffith University School of Medicine, Queensland Health, Brisbane, QLD, Australia, [7] Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia, [8] Institute of Environmental Science and Research, Porirua, New Zealand, [9] School of Medical Sciences, University of Auckland, Auckland, New Zealand, [10] Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, National Infection Service, Public Health England, London, United Kingdom, [11] National Infection Service, Public Health England, Addenbrooke's Hospital, Cambridge, United Kingdom, [12] Scottish MRSA Reference Service, Scottish Microbiology Reference Laboratories, Glasgow Royal Infirmary, Glasgow, United Kingdom, [13] Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal, [14] Laboratory of Microbiology and Infectious Diseases, The Rockefeller University, New York, NY, United States, [15] School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, Australia, [16] Department of Microbiology, Fiona Stanley Hospital, Perth, WA, Australia, [17] Victorian Infectious Disease Service, The Royal Melbourne Hospital, and The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia

Background: In Australia, community-associated methicillin-resistant *Staphylococcus aureus* (MRSA) lineage sequence type (ST) 93 has rapidly risen to dominance since being described in the early 1990s. We examined 459 ST93 genome sequences from Australia, New Zealand, Samoa, and Europe to investigate the evolutionary history of ST93, its emergence in Australia and subsequent spread overseas.

Results: Comparisons with other *S. aureus* genomes indicate that ST93 is an early diverging and recombinant lineage, comprising of segments from the ST59/ST121 lineage and from a divergent but currently unsampled Staphylococcal population. However, within extant ST93 strains limited genetic diversity was apparent with the most recent common ancestor dated to 1977 (95% highest posterior density 1973–1981). An epidemic ST93 population arose from a methicillin-susceptible progenitor in remote Northern Australia, which has a proportionally large Indigenous population, with documented overcrowded housing and a high burden of skin infection. Methicillin-resistance was acquired three times in these regions, with a clade harboring a staphylococcal cassette chromosome *mec* (SCC*mec*) IVa expanding and spreading to Australia's east coast by 2000. We observed sporadic and non-sustained introductions

of ST93-MRSA-IVa to the United Kingdom. In contrast, in New Zealand, ST93-MRSA-IVa was sustainably transmitted with clonal expansion within the Pacific Islander population, who experience similar disadvantages as Australian Indigenous populations. **Conclusion:** ST93 has a highly recombinant genome including portions derived from an early diverging *S. aureus* population. Our findings highlight the need to understand host population factors in the emergence and spread of antimicrobial resistant community pathogens.

INTRODUCTION

Over the past two decades multiple clones of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) have emerged globally (DeLeo et al., 2010). These distinct CA-MRSA lineages have arisen almost simultaneously in different regions of the world. Although the pathogen-specific factors contributing to this phenomenon continue to be sought, a common characteristic is that particular host populations have been identified as risk groups for CA-MRSA infections and/or carriage. These populations tend to be socio-economically marginalized with Indigenous, injecting drug user, incarcerated and homeless populations being overrepresented. For example, USA300 infections in the United States were first reported in children in disadvantaged populations in Chicago (Herold et al., 1998), while CA-MRSA in Australia were initially documented from Indigenous populations in Western Australia (Udo et al., 1993). This association is further supported when examining ST80-MRSA emergence in Europe with infections occurring in children from lower socio-economic households (Dufour et al., 2002).

In contrast to healthcare-associated MRSA lineages, which have demonstrated intercontinental transmission with subsequent establishment and spread within local healthcare systems (Harris et al., 2010; Holden et al., 2013), the typical pattern of movement for CA-MRSA lineages appears to be far more geographically restricted. Despite frequent intercontinental traffic between North America and Europe, the dominant USA300 CA-MRSA lineage has not become established in Europe (Toleman et al., 2016; Bouchiat et al., 2017). Similarly, ST80-MRSA in Europe, which likely arose from a methicillin-susceptible *S. aureus* (MSSA) ancestor from sub-Saharan Africa (Stegger et al., 2014), has not become established in the United States. Frequent intercontinental transfers have also been noted for the South-East Asian CA-MRSA clone (ST59), but with limited evidence of ongoing clonal transmission outside the country of origin (Ward et al., 2016).

In Australia, the dominant CA-MRSA lineage is the Panton–Valentine leukocidin positive ST93. First reported in the early 2000s in Queensland on the eastern coast of Australia (Munckhof et al., 2003), ST93-MRSA has rapidly become the most common circulating CA-MRSA lineage (Coombs et al., 2016). Severe clinical manifestations have been widely reported (Peleg et al., 2005; Tong et al., 2008). Murine models suggest ST93 to be more virulent than other CA-MRSA clones including USA300 (Chua

et al., 2011; Tong et al., 2013). Previous investigations of ST93 genomics by Stinear et al. (2014) demonstrated limited genetic diversity of the clone, with acquisition of staphylococcal cassette chromosome *mec* element (SCC*mec*) conferring methicillin-resistance on two occasions. Dating and phylogeographical analyses concluded that the most recent common ancestor of ST93 MSSA included in that study arose in the 1970s in North Western Australia. These conclusions were based on a collection of 56 isolates predominantly from Western Australia (representing approximately half of all the isolates). Given these limitations, and the current uncertainty of the ancestry of ST93 relative to other *S. aureus* lineages, we examined the genomics of ST93 on an extended dataset (consisting of an additional 403 ST93 isolates). This collection not only represents a deeper sampling from the hypothesized regions of emergence in Australia, but also a broader sampling from non-Australian regions to encompass the presumed international spread of ST93-MRSA. We describe the relationship of ST93 to other *S. aureus* lineages and impact of recombination on its genome, as well as the emergence of ST93 as a successful MSSA and subsequent CA-MRSA lineage from remote sparsely populated regions of Northern Australia.

MATERIALS AND METHODS

Isolate Collection

Overall 459 sequenced *S. aureus* isolates were included in the study. We set out to capture as much diversity as practicable of ST93, and in so doing, established for whole-genome sequencing a collection of 403 presumed ST93 *S. aureus* (based on local typing schemes) from international locations where ST93 has been reported including Australia, New Zealand, Samoa, and five European countries (**Supplementary Table S1**). All isolates were from patients with invasive and non-invasive infections collected between 1991 and 2012 and confirmed as either MSSA or MRSA. The data were supplemented by sequence reads from a further 56 isolates, sequenced as part of a previous study by Stinear et al. (2014). Metadata details can be found in **Supplementary Table S1**.

Sequencing and Assembly

Genomic DNA samples were sequenced on the Illumina HiSeq 2500 platform using 151b paired-end TruSeq chemistry. Reads were mapped to the complete JKD6159 reference

genome (2,811,435 bp, GenBank: NC017338), an ST93-MRSA-IVa isolated from a patient in Melbourne, Australia in 2004 (Chua et al., 2010) using SMALT. Variants were called using a combination of Samtools mpileup[1] and bcftools[2] with filtering for read depth, read base quality, and mapping quality, ensuring only high-quality SNPs were accepted. All variants at either indels or secondary to mobile elements were excluded from the mapping-based analysis.

The Relationship of ST93 Relative to Other *S. aureus* Lineages

Manually completed genomes of *S. aureus* were obtained from the NCTC3000 reference collection[3] using nctc-tools[4] and supplemented with reference sequences from NCBI including the MSHR1132 genome assembly of *Staphylococcus argenteus* (**Supplementary Table S2**). Multi-locus sequence types (MLST) were assigned with *mlst*[5] and GFF files were generated with a standardized re-annotation of the collection with Prokka (Seemann, 2014). The whole-genome alignment consisted of extracted locally collinear blocks (LCBs) of 500 bp, present in all genomes determined by progressiveMauve (Darling et al., 2010) and ordered against the ST93 reference genome. Recombination in the core genome alignment was detected with Gubbins (Croucher et al., 2015). Pre- and postrecombination maximum-likelihood phylogenies (under the GTR+G model) were generated from the SNP matrices [determined using SNP-sites (Page et al., 2016) and from variant sites determined by Gubbins] employing RAxML-NG v.0.5[6] with the *S. argenteus* genome MSHR1132 as outgroup and 100 bootstrap replicates. Resultant phylogenies were visualized in Interactive Tree of Life (Letunic and Bork, 2016). The identity and total number of unique sites affected by recombination was computed by taking the set of base pair locations present in all recombinant segments for a lineage. Recombination events were filtered by size and visualized against the respective non-recombinant phylogeny with *plotTree*[7]. To test the accuracy of the results, the aforementioned workflow was repeated using Mugsy instead of progressiveMauve to establish the core genome alignment (Angiuoli and Salzberg, 2011). Similar results were obtained throughout. Nucleotide divergences were computed on the ProgressiveMauve alignment over a non-overlapping 10,000 bps sliding window under the K80 model of evolution in the *dist.dna* function of *ape* for R.

Phylogenetic Dating

To investigate the temporal evolution of ST93 we used the Bayesian software package BEAST2 (v.2.3) on the 459 ST93 genomes (Bouckaert et al., 2014). A GTR substitution model with gamma correction for among-site variation was implemented

---

[1] http://samtools.sourceforge.net
[2] https://samtools.github.io/bcftools/
[3] http://sanger.ac.uk/resources/downloads/bacteria/nctc/
[4] https://github.com/esteinig/nctc-tools
[5] https://github.com/tseemann/mlst
[6] https://github.com/amkozlov/raxml-ng
[7] https://github.com/katholt/plotTree

using the core genome following masking of homologous recombination identified by Gubbins (Croucher et al., 2015). All combinations of strict, relaxed lognormal, relaxed exponential and random clock models and constant, exponential and skyline population models were evaluated. Models which failed to converge or result in an effective sample size values <200 following 100 million generations, sampling every 1,000 generations, were excluded from further comparison. The best-fit model combination, exponential population with a relaxed lognormal clock model, was determined by marginal likelihoods using the smoothed harmonic mean estimator within the program Tracer v.1.4. A maximum clade credibility tree was created using the treeAnnotator program. Phylogenetic trees were visualized and manipulated with Figtree v1.4.2[8].

Phylogenetic Clustering

To determine clusters in the phylogeny, mutual k-nearest neighbor population graphs (mkNNG) were constructed from the cophenetic distances of the BEAST2 phylogeny using package *ape* for R and NetView (Neuditschko et al., 2012; Steinig et al., 2016) with a range of $k = 1$–100. Fast-greedy modularity optimization (Clauset et al., 2004) Walktrap (Pons and Latapy, 2005) and Infomap (Rosvall and Bergstrom, 2010) algorithms were run on each network configuration to define communities in the graph topology. To assess the stability of the communities in the network topology and select an appropriate configuration, the number of detected communities (n) was plotted against k. The plot indicated a disassociation of the community structure (in line with increasing sparseness of edges) in the network at $k < 10$. Communities found in a conservative and stable configuration of the network ($k = 60$, vertices = 459, edges = 4,711) were selected to determine the general network clusters using the low-resolution fast-greedy modularity optimization algorithm (**Supplementary Figure S7**). Cluster assignments were mapped back to the phylogeny (**Supplementary Figures S7A,B**).

Global and Regional Transmission Routes

A Bayesian phylogeographic analysis was performed within BEAST2 using the longitude and latitude data from where the sequenced isolate originated (Lemey et al., 2009). Subsequent data were visualized and manipulated in the software tool SPREAD (Bielejec et al., 2011). In addition, ML Ancestral state reconstruction was undertaken to confirm the likely origin of ST93 treating isolate location as a discrete trait in R package *ape*. Skyline population models were examined but excluded as they were found to be biased by non-uniform sampling.

SCC*mec* Analysis

Alignment of *de novo* contigs to the reference (JKD6159) SCC*mec*IV[2B] revealed two additional SCC*mec* types. Assembly and closure of a novel SCC*mec* type was achieved by mapping contigs to previously described elements and Sanger sequencing to bridge sequence gaps. The variant was referred

---

[8] http://tree.bio.ed.ac.uk/software/figtree

to the International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (IWG-SCC) which confirmed the new variant, named SCC*mec*IVn. Two "discordant" isolates (from NZ and Victoria) were detected (i.e., *mecA* negative but phenotype reported as methicillin-resistant). Sequence mapping to SCC*mec* confirmed reads mapping to the components of the cassette suggesting these isolates had lost components of their SCC*mec*, a phenomenon known to occur with storage (van Griethuysen et al., 2005).

Pan-Genome and Genomic Association

A pan-genome was created from isolate contigs, which were generated with an in-house assembly and improvement pipeline based on Velvet (Zerbino, 2010) and annotated using Prokka (Seemann, 2014). The pan-genome was generated using Roary (Page et al., 2015) and visualized using Phandango (Hadfield et al., 2017).

Nucleotide Sequence Accession Number

The sequence reads for all isolates are available from ENA (**Supplementary Table S1**).

RESULTS

ST93 Is a Divergent and Recombinant *S. aureus* Lineage

The position of ST93 within the *S. aureus* population structure has never been clearly defined. Indications from multilocus sequence typing loci are that ST93 diverged early from other *S. aureus* but was not ancestral to ST121 or ST59 (Tong et al., 2010). In contrast, when using the entire shared core genome, a maximum likelihood phylogeny places ST93 as the earliest diverging lineage from all other *S. aureus* genomes included in this study, with a bootstrap support of 100% (**Figure 1A**). However, after masking of putative recombined regions in the core genome (21.6% in ST93), ST93 was no longer the earliest diverging *S. aureus* lineage. Instead, it was placed within the broader clade containing ST59 and ST121 (**Figure 1B**). We examined the putative recombinant regions and found no evidence for a large recombined region event, unlike ST239 which has a single chromosomal replacement of ~600 kb (Robinson and Enright, 2004), but rather evidence for numerous short segments with an average length of 2,209 bp (**Supplementary Figures S1, S2**). More than half of the recombinant regions were unique to the ST93 branch and not shared with other lineages (**Supplementary Figures S1, S2B**). Such terminal recombination events were more common and their contribution to nucleotide variation greater in ST93 compared to other *S. aureus* genomes [as measured by the recombination/mutation ratio (r/m) of 1.96 and the number of recombination events to point mutation ratio (ρ/θ) of 0.0167; **Supplementary Figures S2C,D**]. When we examined pair-wise nucleotide divergence of the core genome for the *S. aureus* species we found that for some regions ST93 was equally and highly diverged from the two main branches of *S. aureus* as represented by ST59 and ST8 strains (**Figure 1C**

and **Supplementary Figure S3**). The genomic location of these divergent regions correlated with the regions of terminal recombination events detected by Gubbins and included several complete coding regions (**Supplementary Table S3**). For other regions, the pairwise ST93↔ST59 distance was considerably less than the ST93↔ST8 distance and was similar to that of ST59↔ST121. Thus, ST93 has been heavily impacted by recombination, with large parts of its genome showing a specific relationship with the ST59/ST121 group, and other parts descended from *S. aureus* lineages earlier diverging than any of the genomes in this study.

ST93 Emerged From North-Western Australia Within the Last 50 Years

We determined the phylogenetic structure of a geographically diverse collection of 459 ST93 isolates, which included MSSA ($n = 112$) and MRSA ($n = 347$) (**Figure 2**). The isolates represent an unbiased sampling of *S. aureus* recovered from community and hospital onset skin infections in the Northern Territory (McDonald et al., 2006; Tong et al., 2010), MRSA strains from hospitals and reference collections from Australia, and MRSA strains from public health collections in Europe (mainly the United Kingdom) and New Zealand (**Supplementary Table S1** and **Supplementary Figure S4**). There were 7,044 core genome SNP-sites among the 459 isolates.

Dating of the root of the tree (**Figure 2**) suggests the most recent common ancestor of currently extant ST93 isolates to be 1977 [95% highest posterior density (HPD) 1973–1981; and a mutation rate of 1.05×10^{-6} (95% HPD 9.8×10^{-7} to 1.12×10^{-6})]. The ST93-MSSA isolates form a clear basal group, with the vast majority of MSSA isolates ($n = 104/112$), including the most basal MSSA isolates, originating from the bordering jurisdictions of the Northern Territory and Western Australia (**Figure 2C**). Ancestral state reconstruction strongly supports an origin of ST93-MSSA from North Western Australia (**Figure 3**) with the probability heavily weighted toward a Northern Territory origin rather than Western Australia.

Methicillin Resistance Emerged in ST93 on At Least Three Occasions

Within this diverse MSSA population, SCC*mec* has been acquired on at least three occasions (**Figure 2**). The acquisition of a type IVa SCC*mec* element (SCC*mec*IVa), in the late 1980s or early 1990s, was the basis for a large clonal expansion event, encompassing 332/345 (96.2%) of the MRSA isolates in the study (two isolates were phenotypically MRSA but genotypically lacking *mecA*). A single isolate from Western Australia harbored a type V SCC*mec* element (SCC*mec*V). A separate cluster, consisting of 12 isolates from the Northern Territory was found to harbor a novel SCC*mec* element (GenBank Accession Number KX385846) (**Figure 4**), now designated IVn (SCC*mec*IVn; International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements). The 12 isolates formed a monophyletic cluster, representing a separate and distinct acquisition event, within the Northern Territory. Of the 12 isolates, 11 were *spa* type

FIGURE 1 | Phylogenetic position of ST93 in relation to other *Staphylococcus aureus*. Maximum-likelihood cladograms (RAxML, GTR+G) generated from complete reference genomes from a core alignment with progressiveMauve before **(A)** and after **(B)** removing recombination with Gubbins. Red branches indicate branches with bootstrap support of 100 and indicate strong support for an early divergence of ST93 (indicated by the blue branch) outside the main group of *S. aureus* sequence types **(A)**. Following masking of recombination, ST93 is located within the larger *S. aureus* complex basal to a larger clade containing other regional clones like ST59 and ST121 **(B)**. Colored boxes around tip labels indicate isolate pairs used in subsequent comparisons. **(C)** Pair-wise nucleotide divergence across a 10,000 bp sliding window along the *S. aureus* core genome with comparisons between ST59-ST121, ST93-ST59, and ST93-ST8. Bases affected by recombination as detected by Gubbins across the ST93 genome are represented above with shared ancestral events (red) and events unique to the ST93 terminal branch (blue). Terminal branch recombination regions align with genomic positions where ST93 is highly diverged compared to other *S. aureus*.

t202, which is the majority *spa* type in the entire dataset (346/459) and 1 was *spa* type t1811. These two *spa* types differ by the presence of one repeat (t202, 11-17-23-17-16-16-25; t1811, 11-17-17-16-16-25). SCC*mec*IVn is closely related to SCC*mec*IVa of the reference strain JKD6159 (ST93-MRSA-IVa), and shares >98% nucleotide identity apart from a pUB110 insertion flanked by IS*431* elements in the J3 region of SCC*mec*IVn (**Figure 4**). The pUB110 element displays

>99% identity to pUB110 in the type II SCC*mec* elements (SCC*mec*II) from Mu3, Mu50, MRSA252, and JKD6008, and contains an *ant4-1/aadD* aminoglycoside resistance gene that has been described to confer high-level resistance to tobramycin, amikacin, and kanamycin, as well as the bleomycin resistance gene, *bleO*.

The geographical origins of the ST93-MRSA-IVa lineage have to date remained unclear. ST93-MRSA-IVa was first described
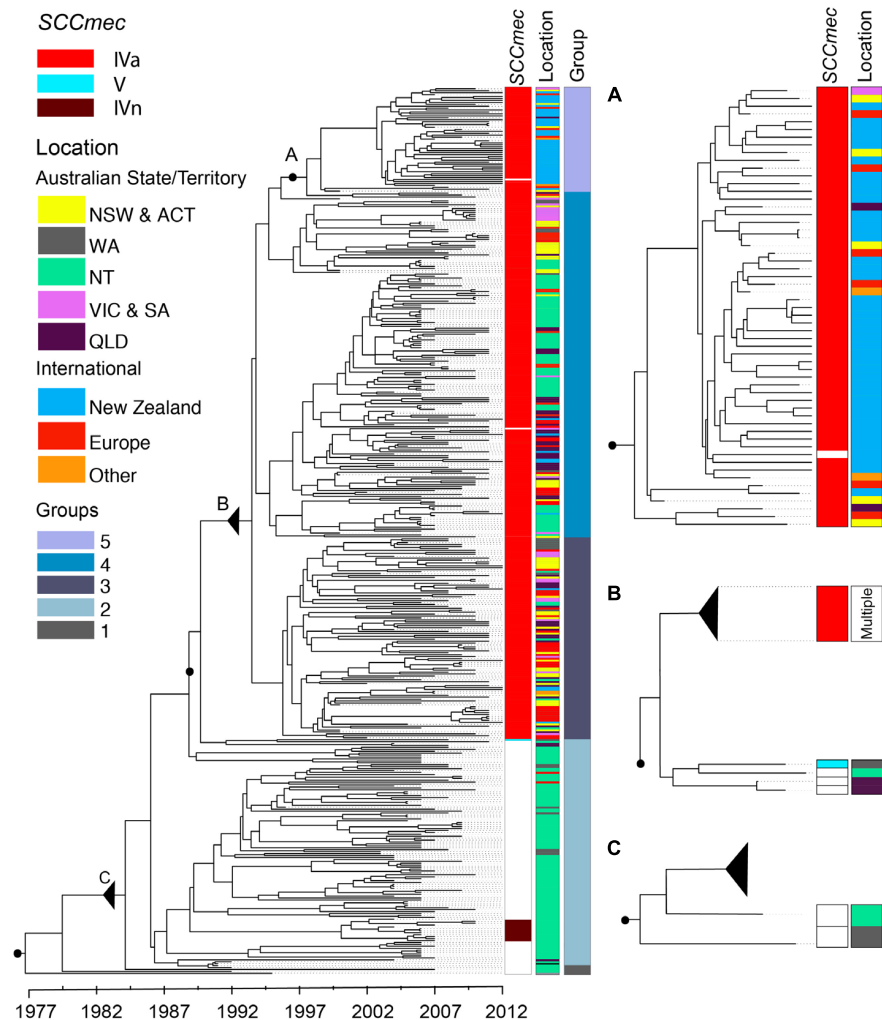
FIGURE 2 | ST93 phylogeny. Maximum clade credibility tree obtained from the core genome not affected by recombination of 459 ST93 sequences using BEAST2. *x*-Axis depicts time in years. Metadata from left to right indicate the presence or absence of the *mecA* gene by SCC*mec* type (i.e., methicillin resistance; white – absence, red – SCC*mec*IVa, brown – SCC*mec*IVn, light blue – SCC*mec*V), the region of origin, and clustering based on the network analysis. Three regions of interest are magnified to the right of the tree with associated metadata. **(A)** Corresponds with group 5 containing predominantly isolates from New Zealand (see text for details). **(B)** Acquisition events of SCC*mec*IVa and SCC*mec*V elements leading to the establishment of ST93-MRSA-IVa. **(C)** The likely origin of ST93 showing the earliest basal strains of ST93-MSSA leading into the main MSSA group. Letters, black filled circles, and solid triangles (collapse point of the rest of the tree) allow for correlation of inserts with the total phylogeny. NZ, New Zealand; EU, Europe; NSW, New South Wales; NT, Northern Territory; QLD, Queensland; VIC, Victoria; WA, Western Australia. Australian Capital Territory (*n* = 1) and South Australian (*n* = 2) isolates are included with NSW and Victorian isolates, respectively.

in South East Queensland in 2000 as a cause of skin and soft tissue infections (SSTI) in mainly Caucasian patients (Munckhof et al., 2003), and then in 2003 as a cause of severe disease in mostly Indigenous Australian patients (Peleg et al., 2005). In our dataset, the earliest ST93-MRSA-IVa strains were isolated in 2000, and eight of these nine strains, were from the eastern states of Australia (New South Wales/Australian Capital Territory: SAPWH23, SAPWH39, SAPRPAH96, SAPTCH92, SAPWH61, SAPWH64; Queensland: SAPRBH98; Victoria: SAPRCH74; and South Australia: SAPGPSA73). Notably, despite comprehensive state-wide typing of MRSA isolates since 1996 in Western Australia, ST93-MRSA was not identified in Western Australia

until 2003. Ancestral state reconstruction (**Figure 3**) indicates that the ST93-MRSA-IVa lineage most likely emerged from the Northern Territory in the early 1990s. By the year 2000, transmission from the Northern Territory to the eastern states of Queensland and New South Wales had occurred.

International Transmission of ST93-MRSA-IVa

Network clustering delineated five main ST93 groups: Groups 1 and 2 representing MSSA strains and groups 3, 4, and 5 representing the three major groups within the ST93-MRSA-IVa clade (**Figure 2**). Group 3 is principally comprised of strains from

FIGURE 3 | ST93 geographic origins. Ancestral state reconstruction based on maximum clade credibility tree generated using BEAST2 with geographic regions as discrete traits. Each node indicates the log likelihood for specified locations as the origin of the node in pie chart format. Pie charts corresponding with the origins of (1) MSSA and (2) MRSA are enlarged for easier visualization. Abbreviations and colors for regions are as per **Figure 2**.

FIGURE 4 | Novel SCC*mec*IVn. Sequence comparison of regions of homology between SCC*mec*II and IV elements from isolates Mu3 (green), ST93 JDK6519 (purple), and the new consensus SCC*mec*IVn from ST93 (blue). Mu3 carries SCC*mec*II and *mec* complex A, JKD6159 SCC*mec*IVa and *mec* complex B, and the novel SCC*mec*IVn is also *mec* complex B. Inner rings denote coding regions (orange – transposition; red – resistance; dark gray – SCC*mec* characteristic regions *orfX*, *mec* cassette and *ccr* complex; light gray – other). Chords denote major regions (*mec*, *ccr*, pUB110) defined by nucleotide BLAST comparisons (>99% identity and >1 kb) of SCC*mec*II and IV elements against the SCC*mec*IVn.

the Australian east coast (81/105 strains are from New South Wales, Australian Capital Territory, Victoria, South Australia, and Queensland) and Europe.

European isolates (mainly from the United Kingdom) are highly diverse and are scattered across groups 3, 4, and 5 consistent with multiple sporadic exportations to the United Kingdom (and possibly elsewhere in Europe). Despite frequent exportation events and a limited number of transmission chains within the United Kingdom [**Supplementary Figure S5**; with one cluster of three isolates originating from the same pediatric hospital (**Supplementary Figure S5** – Box 2) and one cluster of four isolates recovered from an index case and siblings in the same household (**Supplementary Figure S5** – Box 1)], no evidence for sustained transmission in the United Kingdom was seen.

In contrast to the epidemiological picture in Europe, there was evidence for sustained transmission and endemic spread of ST93-MRSA-IVa in New Zealand. The majority of New Zealand isolates cluster together within group 5 (**Figure 2**). This strongly suggests establishment in around 2000 followed by clonal expansion and sustained transmission of a New Zealand specific ST93-MRSA-IVa lineage. In addition, not only is there evidence of ongoing transmission in New Zealand, but also instances of transmission back to Australia and to Europe. Given the novelty of this observation with respect to CA-MRSA lineages (i.e., introduction and establishment of a CA-MRSA lineage), we examined the epidemiological characteristics associated with these New Zealand isolates. Significantly, 52% (28/54) of all ST93-MRSA-IVa from New Zealand [or 63% (25/40) of group 5] were centered around Auckland on the north island and principally originating from within the Māori and Pacific Islander populations [of New Zealand isolates with a known host ethnicity, 32/49 (65%) of ST93-MRSA-IVa overall, and 24/38 (63%) of group 5, were recovered from Māori and Pacific Islanders]. A Spatial Phylogenetic reconstruction (**Figure 5**) supported transmissions between Australia and New Zealand via hosts of European or unknown ethnicity, while the majority of within New Zealand transmission was within the Pacific Islander population.

The contrast between the epidemiology in the United Kingdom and New Zealand is further supported by the longitudinal data (**Supplementary Figure S6**) based on contemporary MRSA typing results (between 2008 and 2015) from the United Kingdom, New Zealand and Australia, which indicate consistently low levels of ST93 in the United Kingdom and therefore likely ongoing sporadic importation, compared to continued increases in Australia and New Zealand consistent with sustained transmission and endemic spread within these two countries.

Genomic Markers of Specific Clades

We examined the dataset for genetic markers that may provide clues to the success of specific clades within the overall ST93 population. Twelve single nucleotide variants (SNVs) mapped to the branch corresponding to SCC*mec*IVa acquisition, of which 11 SNVs were not associated with known virulence genes. One SNV introduced a premature stop codon in *aryK*, a novel virulence regulator as previously described by Chua et al. (2014). It has been hypothesized that this loss-of-function point mutation may act to reduce the virulence of ST93 and

FIGURE 5 | Transmission of ST93-MRSA-IVa in Australia and New Zealand. Phylogenetic diffusion of ST93 within New Zealand based on discrete traits implemented using SpreaD3. Oceania region is depicted with the outlined area (New Zealand) enlarged in insert B. White open circles represent the origin of the isolate. For isolates from the same location a jiggle factor is included such that intra-local diffusion is visible. Connecting colored lines represent the ethnicity of the host as shown in the legend. Circular polygons are proportional to the number of tree lineages maintaining that location. Indigenous ethnicity reflects either an Aboriginal Torres Strait Islander (Australian) or Maori (New Zealand) host. The analysis shows that within New Zealand maintenance of ST93 is centered on Auckland within the Maori and Pacific Islander communities and that introductions of this clone are linked to travel between New Zealand and Australia.

contribute to host persistence. However, the ongoing circulation of ST93-MSSA in the Northern Territory suggests that ST93 is quite capable of successful persistence despite the lack of an SCC*mec* element and the absence of this loss-of-function point mutation.

We also looked at variants restricted to the New Zealand clade (group 5) (**Supplementary Table S4**). Fourteen SNVs were found of which eight were non-synonymous. Of note, a mutation in *adsA* (also known as *sasH*), encoding a surface bound protein adenosine synthase, was detected. *AdsA* is a critical virulence factor in *S. aureus* and converts adenosine tri-, di-, and monophosphate to adenosine, facilitating escape from phagocytic clearance (Lemey et al., 2009; Bielejec et al., 2011). However, the P33S substitution conferred by the mutation sits outside both the metallophosphatase and 5′-nucleotidase conserved domains of *adsA*, and thus the functional implications of this mutation with regards to fitness or virulence are unknown, and will require further laboratory experiments to elucidate.

The overall ST93 gene content (pan-genome) consisted of 3,837 genes, of which the core genome (present in over 99% of isolates) comprised 2,436 genes. Of the remaining 1,401 genes outside of the core genome, 1,212 were found in <15% of isolates, consistent with a significant flux of accessory genes. Apart from genes associated with SCC*mec* elements, we did not find associations between particular accessory genome elements

and the major clades. 455/459 isolates harbored the PVL genes *lukS-PV* and *lukF-PV*. The Roary output for the pan-genome can be found in **Supplementary Table S5**.

DISCUSSION

The evolutionary history of ST93 represents the emergence from isolated and sparsely populated regions in Australia of a virulent community-associated *S. aureus* lineage, which has become established more broadly in Australia. Global transmission has also occurred in different patterns, with sporadic transmission to the United Kingdom and elsewhere in Europe and contrastingly, sustained endemicity in the New Zealand Pacific Island population. These findings suggest that as much as pathogen characteristics are important, aspects of the host population play a significant role in the success of bacterial lineages.

At a whole genome level, ST93 can be considered the earliest diverging of currently sampled *S. aureus* lineages. We found evidence that recombination has played a significant role in the evolutionary history of ST93 with large proportions of the genome derived from highly divergent sequences and other segments from ancestors related to the ST59/ST121 clade of *S. aureus*. Somewhat at odds with this ancient history is that the currently extant ST93 population has limited genetic diversity with a most recent common ancestor from the 1970s.

The current ST93 population has rapidly become established in geographically disparate remote regions of the Northern Territory and Western Australia with largely Indigenous Australian populations. In a survey of Indigenous Australian residents from remote communities in northern Western Australia from 1996 to 2003, the most common lineage isolated was ST93-MSSA (O'Brien et al., 2009). Broader epidemiological studies with an unbiased sampling of both MSSA and MRSA in Australia concur with a substantial presence of ST93-MSSA in the Northern Territory (Tong et al., 2010) and Western Australia (O'Brien et al., 2009) and its near absence in other parts of Australia (Holmes et al., 2014). A possible interpretation of our results is that ST93, with genomic regions derived from an early diverging population of *S. aureus*, has been carried in isolation by Indigenous Australians, also an early diverging population within *Homo sapiens* (Malaspinas et al., 2016). However, the dating of the emergence of ST93 to the 1970s is inconsistent with this model unless there was a tight bottleneck in the 1970s leading to a marked reduction in diversity of ST93. An alternative scenario is that the expansion of ST93 in humans followed a cross-species transfer event from a yet to be identified non-human reservoir.

The recent expansion and establishment of a *S. aureus* clone may be somewhat surprising given the overall sparseness of population in these regions; the Northern Territory has a population density of 0.2 person per km^2, and the Kimberley region in the north of Western Australia <0.1 person per km^2, compared to 2.9 and 269 people per km^2 for Australia and the United Kingdom, respectively. Notably, the Northern Territory and Kimberley, Western Australia have among the highest proportions of Indigenous people as part of the population in Australia at 26% (and 51% if excluding the urban

center of Darwin) and 41%, respectively. There is documented overcrowding within remote community households with a median of 7 and up to 23 residents per household (Vino et al., 2017) and a high prevalence of SSTI (an estimated 45% of remote living Indigenous children from the Northern Territory, Western Australia and Queensland suffer from impetigo at any one point in time) (Bowen et al., 2015), facilitating the spread and amplification of local emerging *S. aureus* clones. Studies in the United States and United Kingdom have highlighted the importance of household transmission and markers of socio-economic deprivation to the transmission of CA-MRSA lineages (Knox et al., 2015; Tosas Auguet et al., 2016).

Strikingly, ST93-MRSA-IVa was not detected in Western Australia until 2003. However, since then ST93-MRSA-IVa has become the dominant CA-MRSA in Western Australia, with a focus in the north-west of the state, particularly the Kimberley region where ST93-MRSA has been isolated at a rate of 2,499 isolations per 100,000 population in 2015–2016 (compared to 41 isolations per 100,000 population in the metropolitan region around the capital Perth) (Coombs et al., 2017). As with the emergence of ST93-MSSA, it is likely that socio-demographic factors play an important role in this rapid expansion of ST93-MRSA-IVa.

Following its emergence in north-western Australia, ST93-MRSA-IVa became endemic in Australia, and then was transmitted from the Australian eastern seaboard to the United Kingdom and New Zealand. Similar to patterns seen with USA300 (Bouchiat et al., 2017) and ST80 (Stegger et al., 2014), inter-country transmission has not led to continued transmission and endemicity of ST93-MRSA-IVa in the United Kingdom. In contrast, we have demonstrated establishment of ST93-MRSA-IVa in New Zealand, centered primarily in the Auckland region and in the Pacific Islander population. Māori and Pacific Islanders made up 65% of New Zealand ST93-MRSA-IVa infections and are thus clearly over-represented in light of these population groups representing 15 and 7%, respectively of the general New Zealand population. Notably, rates of *S. aureus* skin infections and CA-MRSA infections are significantly higher in Mâori and Pacific Islander populations in New Zealand compared to patients of European ethnicity (Williamson et al., 2013, 2014). The reasons for this association are multifactorial, but are likely to include a combination of poverty and domestic overcrowding – i.e., similar socio-demographic conditions to those favoring the emergence of ST93 in north-western Australia (Williamson et al., 2013, 2014). These data may explain why other CA-MRSA lineages remain currently geographically restricted, despite numerous intercontinental transfer events, as incursion events are not usually occurring into a vulnerable population.

CONCLUSION

ST93 is an early diverging lineage of currently sampled *S. aureus* lineages, with a genome shaped by recombination and of which large proportions are derived from highly divergent sequences. The evolutionary history of ST93 suggests the importance of host population characteristics on both the emergence and ongoing spread of CA-MRSA lineages. Understanding this interplay better would enable interventions at controlling these clones to be refined.

DATA AVAILABILITY

The sequence reads for all isolates are available from ENA (**Supplementary Table S1**).

AUTHOR CONTRIBUTIONS

SH, ES, PG, and ST designed and undertook the analysis. SH, PG, and ST wrote the manuscript. PA conducted aspects of the bioinformatic analyses. SRH and MH provided bioinformatic advice and interpreted the results. SH, PG, DH, and ST designed the study. SB, JP, PG, and ST were responsible for the management of the study. GN, DW, HH, SR, AK, ME, ED, HL, GC, DH, PG, and ST identified and collected isolates and relevant clinical information. All of the authors read, modified, and approved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01453/full#supplementary-material

FIGURE S1 | Recombinant regions across *Staphylococcus aureus* as detected by Gubbins. Maximum-likelihood cladograms (RAxML, GTR+G) of the progressiveMauve (2,157,328 sites, 358,325 SNPs) core alignment after removing recombinant segments (blocks). Recombinant segments were filtered by length as indicated and represent shared ancestral events (red) and events unique to terminal branches (blue). Large segments spanning the origin of replication are present in the ST239/ST240 lineage, while smaller events are particularly common in *S. argenteus* and lineages ST22, ST93, ST152, and ST395.

FIGURE S2 | Terminal recombination events. Terminal recombination events (unique to each sequence type) detected with Gubbins in the progressiveMauve core genome alignment (2,157,328 sites). **(A)** Length distribution of unique recombinant events demonstrating a majority of short recombination events (average: 2,209 base pairs) detected in ST93. **(B)** Total base pairs of non-overlapping recombinant events ancestrally shared (red) or unique (blue) across sequence types in our data set. ST93 (top) features the largest number of unique recombinant sites compared to other sequence types. **(C)** Ratio (r/m) of base substitutions predicted to have been imported through recombination (r) to those occurring through point mutations (m), showing only sequence types with

unique (terminal) recombination events on their respective branch of the phylogeny inferred with Gubbins. Thus, there is a strong influence of recombination on the variation accumulated in ST93. **(D)** Ratio (rho/theta) of the number of recombination events (rho) to point mutations, showing a comparatively high rate of recombination in the core genome relative to mutation for ST93.

FIGURE S3 | Additional pair-wise nucleotide divergence comparisons between representative *S. aureus* genomes including ST93. Pair-wise nucleotide divergence across a 10,000 bp sliding window along the *S. aureus* core genome with comparisons between ST59-ST121, ST93-ST59, ST93-ST8, ST59-ST8, and ST8-ST5. Bases affected by recombination as detected by Gubbins across the ST93 genome are represented above with shared ancestral events (red) and events unique to the ST93 terminal branch (blue). Terminal branch recombination regions align with genomic positions where ST93 is highly diverged compared to other *S. aureus*.

FIGURE S4 | Source of isolates. **(A)** Depicts the location from where the isolates originated with red boxes representing states and territories of Australia. European countries include England (*n* = 45), Scotland (*n* = 12), Denmark (*n* = 4), France (*n* = 1), and Italy (*n* = 1). Note, Samoan isolates (*n* = 4) are not depicted. The number of MSSA isolates sequenced, by location are shown on the secondary axis by the blue line. **(B)** Represents the temporal distribution of the isolates with the number of blood stream isolates depicted by the red bars on the secondary axis. Note that data on source of infection was unknown for 52 isolates.

FIGURE S5 | Evidence of short transmission chains in the United Kingdom. Maximum clade credibility tree as in **Figure 1**, but with transmission chains from the United Kingdom highlighted. Box 1 – four isolates from three individuals of which two were siblings. Box 2 – three isolates from separate patients in same pediatric hospital. Box 3 – there were no obvious epidemiological links with cases

geographically widely distributed. Box 4 – there were no obvious epidemiological links.

FIGURE S6 | Proportion of methicillin-resistant *Staphylococcus aureus* that are ST93-MRSA. This represents longitudinal data of ongoing MRSA typing data from Australia, New Zealand (NZ) and United Kingdom (UK). Graph depicts the percentage of MRSA isolates that are ST93 between 2008 and 2015. Data were obtained from Public Health England, United Kingdom, the Institute of Environmental Science and Research, New Zealand, and the Australian Collaborating Centre for Enterococcus and *Staphylococcus* Species (ACCESS) Typing and Research, Australia.

FIGURE S7 | Network clusters of ST93. **(A)** General network clusters using the low-resolution fast-greedy modularity optimization algorithm. **(B)** Cluster assignments were mapped back to the phylogeny.

TABLE S1 | Details of 459 ST93 *Staphylococcus aureus* genomes.

TABLE S2 | Details of *Staphylococcus aureus* genomes used to determine phylogenetic position of ST93.

TABLE S3 | Coding sequences in highly divergent regions of ST93. The divergence peaks in **Figure 1** were determined using a sliding window of 10,000 bp to assess divergence between the pairwise combinations of sequence types. These annotated genes are extracted from each plateau (i.e., 10,000 bp of divergent peak) and from pairwise combinations involving ST93 where >0.05 divergence was noted.

TABLE S4 | Non-synonymous mutations restricted to the branch defining the New Zealand clade.

TABLE S5 | Pan-genome details from Roary output.

REFERENCES

Angiuoli, S. V., and Salzberg, S. L. (2011). Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* 27, 334–342. doi: 10.1093/bioinformatics/btq665

Bielejec, F., Rambaut, A., Suchard, M. A., and Lemey, P. (2011). SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* 27, 2910–2912. doi: 10.1093/bioinformatics/btr481

Bouchiat, C., Curtis, S., Spiliopoulou, I., Bes, M., Cocuzza, C., Codita, I., et al. (2017). MRSA infections among patients in the emergency department: a European multicentre study. *J. Antimicrob. Chemother.* 72, 372–375. doi: 10.1093/jac/dkw431

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C. H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Bowen, A. C., Mahe, A., Hay, R. J., Andrews, R. M., Steer, A. C., Tong, S. Y., et al. (2015). The global epidemiology of impetigo: a systematic review of the population prevalence of impetigo and pyoderma. *PLoS One* 10:e0136789. doi: 10.1371/journal.pone.0136789

Chua, K., Seemann, T., Harrison, P. F., Davies, J. K., Coutts, S. J., Chen, H., et al. (2010). Complete genome sequence of *Staphylococcus aureus* strain JKD6159, a unique Australian clone of ST93-IV community methicillin-resistant *Staphylococcus aureus*. *J. Bacteriol.* 192, 5556–5557. doi: 10.1128/JB.00878-10

Chua, K. Y., Monk, I. R., Lin, Y. H., Seemann, T., Tuck, K. L., Porter, J. L., et al. (2014). Hyperexpression of alpha-hemolysin explains enhanced virulence of sequence type 93 community-associated methicillin-resistant *Staphylococcus aureus*. *BMC Microbiol.* 14:31. doi: 10.1186/1471-2180-14-31

Chua, K. Y., Seemann, T., Harrison, P. F., Monagle, S., Korman, T. M., Johnson, P. D., et al. (2011). The dominant Australian community-acquired methicillin-resistant *Staphylococcus aureus* clone ST93-IV [2B] is highly virulent and genetically distinct. *PLoS One* 6:e25887. doi: 10.1371/journal.pone.0025887

Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 70(6 Pt 2):066111. doi: 10.1103/PhysRevE.70.066111

Coombs, G., Pearson, J., and Robinson, J. O. (2017). *Western Australian Methicillin-Resistant Staphylococcus aureus Epidemiology and Typing Report July 1 2016 to June 30 2017*. Australian Collaborating Centre for Enterococcus and Staphylococcus Species (Access) Typing and Research. Available at: https://ww2.health.wa.gov.au/~/media/Files/Corporate/general%20documents/Infectious%20diseases/PDF/HISWA/Annual%20reports/WA_annual_report_MRSA_2016_2017.pdf

Coombs, G. W., Daley, D. A., Thin Lee, Y., Pearson, J. C., Robinson, J. O., Nimmo, G. R., et al. (2016). Australian group on antimicrobial resistance Australian *Staphylococcus aureus* sepsis outcome programme annual report, 2014. *Commun. Dis. Intell. Q. Rep.* 40, E244–E254.

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43:e15. doi: 10.1093/nar/gku1196

Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147

DeLeo, F. R., Otto, M., Kreiswirth, B. N., and Chambers, H. F. (2010). Community-associated meticillin-resistant *Staphylococcus aureus*. *Lancet* 375, 1557–1568. doi: 10.1016/S0140-6736(09)61999-1

Dufour, P., Gillet, Y., Bes, M., Lina, G., Vandenesch, F., Floret, D., et al. (2002). Community-acquired methicillin-resistant *Staphylococcus aureus* infections in France: emergence of a single clone that produces Panton-Valentine leukocidin. *Clin. Infect. Dis.* 35, 819–824. doi: 10.1086/342576

Hadfield, J., Croucher, N. J., Goater, R. J., Abudahab, K., Aanensen, D. M., and Harris, S. R. (2017). Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* doi: 10.1093/bioinformatics/btx610 [Epub ahead of print].

Harris, S. R., Feil, E. J., Holden, M. T., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395

Herold, B. C., Immergluck, L. C., Maranan, M. C., Lauderdale, D. S., Gaskin, R. E., Boyle-Vavra, S., et al. (1998). Community-acquired methicillin-resistant

Staphylococcus aureus in children with no identified predisposing risk. *JAMA* 279, 593–598. doi: 10.1001/jama.279.8.593

Holden, M. T., Hsu, L. Y., Kurt, K., Weinert, L. A., Mather, A. E., Harris, S. R., et al. (2013). A genomic portrait of the emergence, evolution and global spread of a methicillin resistant *Staphylococcus aureus* pandemic. *Genome Res.* 23, 653–664. doi: 10.1101/gr.147710.112

Holmes, N. E., Turnidge, J. D., Munckhof, W. J., Robinson, J. O., Korman, T. M., O'Sullivan, M. V., et al. (2014). Genetic and molecular predictors of high vancomycin MIC in *Staphylococcus aureus* bacteremia isolates. *J. Clin. Microbiol.* 52, 3384–3393. doi: 10.1128/JCM.01320-14

Knox, J., Uhlemann, A. C., and Lowy, F. D. (2015). *Staphylococcus aureus* infections: transmission within households and the community. *Trends Microbiol.* 23, 437–444. doi: 10.1016/j.tim.2015.03.007

Lemey, P., Rambaut, A., Drummond, A. J., and Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5:e1000520. doi: 10.1371/journal.pcbi.1000520

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Malaspinas, A. S., Westaway, M. C., Muller, C., Sousa, V. C., Lao, O., Alves, I., et al. (2016). A genomic history of Aboriginal Australia. *Nature* 538, 207–214. doi: 10.1038/nature18299

McDonald, M., Dougall, A., Holt, D., Huygens, F., Oppedisano, F., Giffard, P. M., et al. (2006). Use of a single-nucleotide polymorphism genotyping system to demonstrate the unique epidemiology of methicillin-resistant *Staphylococcus aureus* in remote Aboriginal communities. *J. Clin. Microbiol.* 44, 3720–3727. doi: 10.1128/JCM.00836-06

Munckhof, W. J., Schooneveldt, J., Coombs, G. W., Hoare, J., and Nimmo, G. R. (2003). Emergence of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) infection in Queensland, Australia. *Int. J. Infect. Dis.* 7, 259–267. doi: 10.1016/S1201-9712(03)90104-4

Neuditschko, M., Khatkar, M. S., and Raadsma, H. W. (2012). NetView: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. *PLoS One* 7:e48375. doi: 10.1371/journal.pone.0048375

O'Brien, F. G., Coombs, G. W., Pearman, J. W., Gracey, M., Moss, F., Christiansen, K. J., et al. (2009). Population dynamics of methicillin-susceptible and -resistant *Staphylococcus aureus* in remote communities. *J. Antimicrob. Chemother.* 64, 684–693. doi: 10.1093/jac/dkp285

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., et al. (2016). SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial. Genomics* 2:e000056. doi: 10.1099/mgen.0.000056

Peleg, A. Y., Munckhof, W. J., Kleinschmidt, S. L., Stephens, A. J., and Huygens, F. (2005). Life-threatening community-acquired methicillin-resistant *Staphylococcus aureus* infection in Australia. *Eur. J. Clin. Microbiol. Infect. Dis.* 24, 384–387. doi: 10.1007/s10096-005-1329-3

Pons, P., and Latapy, M. (2005). Computing communities in large networks using random walks. *J. Graph Algorithms Appl.* 10, 191–218. doi: 10.7155/jgaa.00124

Robinson, D. A., and Enright, M. C. (2004). Evolution of *Staphylococcus aureus* by large chromosomal replacements. *J. Bacteriol.* 186, 1060–1064. doi: 10.1128/JB.186.4.1060-1064.2004

Rosvall, M., and Bergstrom, C. T. (2010). Mapping change in large networks. *PLoS One* 5:e8694. doi: 10.1371/journal.pone.0008694

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Stegger, M., Wirth, T., Andersen, P. S., Skov, R. L., De Grassi, A., Simoes, P. M., et al. (2014). Origin and evolution of European community-acquired methicillin-resistant *Staphylococcus aureus*. *mBio* 5:e1044-14. doi: 10.1128/mBio.01044-14

Steinig, E. J., Neuditschko, M., Khatkar, M. S., Raadsma, H. W., and Zenger, K. R. (2016). Netview p: a network visualization tool to unravel complex population structure using genome-wide SNPs. *Mol. Ecol. Resour.* 16, 216–227. doi: 10.1111/1755-0998.12442

Stinear, T. P., Holt, K. E., Chua, K., Stepnell, J., Tuck, K. L., Coombs, G., et al. (2014). Adaptive change inferred from genomic population analysis of the ST93

epidemic clone of community-associated methicillin-resistant *Staphylococcus aureus*. *Genome Biol. Evol.* 6, 366–378. doi: 10.1093/gbe/evu022

Toleman, M. S., Reuter, S., Coll, F., Harrison, E. M., Blane, B., Brown, N. M., et al. (2016). Systematic surveillance detects multiple silent introductions and household transmission of methicillin-resistant *Staphylococcus aureus* USA300 in the East of England. *J. Infect. Dis.* 214, 447–453. doi: 10.1093/infdis/jiw166

Tong, S. Y., Anstey, N. M., Lum, G. D., Lilliebridge, R. A., Stephens, D. P., and Currie, B. J. (2008). Fatal community-associated methicillin-resistant *Staphylococcus aureus* pneumonia after influenza. *Med. J. Aust.* 188:61.

Tong, S. Y., Lilliebridge, R. A., Bishop, E. J., Cheng, A. C., Holt, D. C., McDonald, M. I., et al. (2010). Clinical correlates of Panton-Valentine leukocidin (PVL), PVL isoforms, and clonal complex in the *Staphylococcus aureus* population of Northern Australia. *J. Infect. Dis.* 202, 760–769. doi: 10.1086/655396

Tong, S. Y., Sharma-Kuinkel, B. K., Thaden, J. T., Whitney, A. R., Yang, S. J., Mishra, N. N., et al. (2013). Virulence of endemic nonpigmented northern Australian *Staphylococcus aureus* clone (clonal complex 75, *S. argenteus*) is not augmented by staphyloxanthin. *J. Infect. Dis.* 208, 520–527. doi: 10.1093/infdis/jit173

Tosas Auguet, O., Betley, J. R., Stabler, R. A., Patel, A., Ioannou, A., Marbach, H., et al. (2016). Evidence for community transmission of community-associated but not health-care-associated methicillin-resistant *Staphylococcus Aureus* strains linked to social and material deprivation: spatial analysis of cross-sectional data. *PLoS Med.* 13:e1001944. doi: 10.1371/journal.pmed.1001944

Udo, E. E., Pearman, J. W., and Grubb, W. B. (1993). Genetic analysis of community isolates of methicillin-resistant *Staphylococcus aureus* in Western Australia. *J. Hosp. Infect.* 25, 97–108. doi: 10.1016/0195-6701(93)90100-E

van Griethuysen, A., van Loo, I., van Belkum, A., Vandenbroucke-Grauls, C., Wannet, W., van Keulen, P., et al. (2005). Loss of the mecA gene during storage of methicillin-resistant *Staphylococcus aureus* strains. *J. Clin. Microbiol.* 43, 1361–1365. doi: 10.1128/JCM.43.3.1361-1365.2005

Vino, T., Singh, G. R., Davison, B., Campbell, P. T., Lydeamore, M. J., Robinson, A., et al. (2017). Indigenous Australian household structure: a simple data collection tool and implications for close contact transmission of communicable diseases. *PeerJ* 5:e3958. doi: 10.7717/peerj.3958

Ward, M. J., Goncheva, M., Richardson, E., McAdam, P. R., Raftis, E., Kearns, A., et al. (2016). Identification of source and sink populations for the emergence and global spread of the East-Asia clone of community-associated MRSA. *Genome Biol.* 17:160. doi: 10.1186/s13059-016-1022-0

Williamson, D. A., Ritchie, S. R., Lennon, D., Roberts, S. A., Stewart, J., Thomas, M. G., et al. (2013). Increasing incidence and sociodemographic variation in community-onset *Staphylococcus aureus* skin and soft tissue infections in New Zealand children. *Pediatr. Infect. Dis. J.* 32, 923–925. doi: 10.1097/INF.0b013e3182905f3d

Williamson, D. A., Zhang, J., Ritchie, S. R., Roberts, S. A., Fraser, J. D., and Baker, M. G. (2014). *Staphylococcus aureus* infections in New Zealand, 2000-2011. *Emerg. Infect. Dis.* 20, 1156–1161. doi: 10.3201/eid2007.131923

Zerbino, D. R. (2010). Using the Velvet de novo assembler for short-read sequencing technologies. *Curr. Protoc. Bioinformatics* 31, 11.5.1–11.5.12. doi: 10.1002/0471250953.bi1105s31

Phylogenomic Classification and the Evolution of Clonal Complex 5 Methicillin-Resistant *Staphylococcus aureus* in the Western Hemisphere

Lavanya Challagundla[1], Jinnethe Reyes[2], Iftekhar Rafiqullah[3], Daniel O. Sordelli[4], Gabriela Echaniz-Aviles[5], Maria E. Velazquez-Meza[5], Santiago Castillo-Ramírez[6], Nahuel Fittipaldi[7,8,9], Michael Feldgarden[10], Sinéad B. Chapman[11], Michael S. Calderwood[12], Lina P. Carvajal[2], Sandra Rincon[2], Blake Hanson[13,14], Paul J. Planet[15], Cesar A. Arias[2,13,14], Lorena Diaz[2] and D. Ashley Robinson[3]*

[1] Department of Data Science, University of Mississippi Medical Center, Jackson, MS, United States, [2] Molecular Genetics and Antimicrobial Resistance Unit, International Center for Microbial Genomics, Universidad El Bosque, Bogota, Colombia, [3] Department of Microbiology and Immunology, University of Mississippi Medical Center, Jackson, MS, United States, [4] Instituto de Investigaciones en Microbiología y Parasitología Médica, Universidad de Buenos Aires and Consejo Nacional de Investigaciones Ciencias y Tecnicas, Buenos Aires, Argentina, [5] Department of Vaccine Evaluation, Instituto Nacional de Salud Pública, Cuernavaca, Mexico, [6] Programa de Genómica Evolutiva, Centro de Ciencias Génomicas, Universidad Nacional Autónoma de México, Cuernavaca, Mexico, [7] Public Health Ontario Laboratory, Toronto, ON, Canada, [8] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada, [9] Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada, [10] National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, United States, [11] Broad Institute of MIT and Harvard, Cambridge, MA, United States, [12] Section of Infectious Disease and International Health, Dartmouth–Hitchcock Medical Center, Lebanon, NH, United States, [13] Division of Infectious Diseases and Center for Antimicrobial Resistance and Microbial Genomics, University of Texas Health Science Center, McGovern Medical School, Houston, TX, United States, [14] Center for Infectious Diseases, School of Public Health, University of Texas Health Science Center, Houston, TX, United States, [15] Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, United States

Clonal complex 5 methicillin-resistant *Staphylococcus aureus* (CC5-MRSA) includes multiple prevalent clones that cause hospital-associated infections in the Western Hemisphere. Here, we present a phylogenomic study of these MRSA to reveal their phylogeny, spatial and temporal population structure, and the evolution of selected traits. We studied 598 genome sequences, including 409 newly generated sequences, from 11 countries in Central, North, and South America, and references from Asia and Europe. An early-branching CC5-Basal clade is well-dispersed geographically, is methicillin-susceptible and MRSA predominantly of ST5-IV such as the USA800 clone, and includes separate subclades for avian and porcine strains. In the early 1970s and early 1960s, respectively, two clades appeared that subsequently underwent major expansions in the Western Hemisphere: a CC5-I clade in South America and a CC5-II clade largely in Central and North America. The CC5-I clade includes the ST5-I Chilean/Cordobes clone, and the ST228-I South German clone as an early offshoot, but is distinct from other ST5-I clones from Europe that nest within CC5-Basal. The CC5-II clade includes divergent strains of the ST5-II USA100 clone, various other clones, and most known vancomycin-resistant strains of *S. aureus*, but is distinct from ST5-II strain N315 from Japan that nests within CC5-Basal. The recombination rate of CC5 was much lower than has been reported for other *S. aureus* genetic backgrounds, which indicates that recurrence of vancomycin resistance in CC5 is not likely due to an enhanced promiscuity. An increased number of antibiotic resistances

and decreased number of toxins with distance from the CC5 tree root were observed. Of note, the expansions of the CC5-I and CC5-II clades in the Western Hemisphere were preceded by convergent gains of resistance to fluoroquinolone, macrolide, and lincosamide antibiotics, and convergent losses of the staphylococcal enterotoxin p (*sep*) gene from the immune evasion gene cluster of phage φSa3. Unique losses of surface proteins were also noted for these two clades. In summary, our study has determined the relationships of different clades and clones of CC5 and has revealed genomic changes for increased antibiotic resistance and decreased virulence associated with the expansions of these MRSA in the Western Hemisphere.

INTRODUCTION

Methicillin-resistant *Staphylococcus aureus* (MRSA) is among the leading causes of antibiotic-resistant bacterial infections in hospital and community settings (Centers for Disease Control and Prevention, 2013; World Health Organization, 2014). These infections can range in severity from superficial skin infections to life-threatening invasive infections such as sepsis, infective endocarditis, and osteomyelitis (Boucher et al., 2010). Different MRSA clones predominate in different geographic regions and can differ in the rates and types of infections that they cause (Grundmann et al., 2010; David et al., 2014). Over the past 20 years, MRSA related to multilocus sequence type (ST) 5, which is classified into clonal complex (CC) 5, have been among the most prevalent clones causing hospital-associated infections in the Western Hemisphere (Aires De Sousa et al., 2001; Christianson et al., 2007; Rodriguez-Noriega et al., 2010; Diekema et al., 2014; Arias et al., 2017; Tickler et al., 2017). CC5-MRSA clones can differ in their Staphylococcal Chromosomal Cassette *mec* (SCC*mec*) genetic element, which carries the resistance determinant to anti-staphylococcal β-lactams, and in gene content (Christianson et al., 2007; Monecke et al., 2011). The identification of MRSA clones based on ST-SCC*mec* types (Robinson and Enright, 2004) has allowed for more precise transnational communication and tracking of these clones, and characterization of gene content has provided additional markers and insights into the lifestyles of these clones (Monecke et al., 2011).

Studies have identified the ST5-I Chilean/Cordobes clone (Sola et al., 2002), the ST5-II Canadian MRSA-2, New York/Japan, and USA100 clones (Roberts et al., 1998; McDougal et al., 2003; Christianson et al., 2007), and the ST5-IV USA800 clone (McDougal et al., 2003), as particularly common in the Western Hemisphere. For example, in the United States, ST5-II may have accounted for >40% of MRSA infections since the mid-1990s (Roberts et al., 1998; McDougal et al., 2003). In the southern cone of South America, ST5-I may have accounted for >60% of MRSA infections since the early 2000s (Sola et al., 2006). More recently in this region of South America, ST5-IV with *pvl* and *sea* toxin genes have

emerged as a cause of community-associated infections and hospital infections especially among children (Sola et al., 2008; Egea et al., 2014). Other CC5-MRSA clones with SCC*mec* types I–VII have been reported elsewhere in the world (Monecke et al., 2011). In some cases, such as the ST225-II Rhine-Hesse clone in Central Europe (Nubel et al., 2010) and ST5-II in Western Australia (Coombs et al., 2007), the Western Hemisphere has been implicated as the origin of the clones.

Importantly, CC5 is the principal genetic background within *S. aureus* upon which full resistance to vancomycin—one of the drugs of choice for treating MRSA infections—repeatedly has arisen by acquisition of the *vanA* operon from *Enterococcus* spp. (Kos et al., 2012). The reasons why CC5 and not other *S. aureus* backgrounds has been the focal point for these acquisitions are not clear, but could include increased promiscuity, a genome that is conducive to *vanA* expression, or a shared niche with *Enterococcus* spp. In addition to the substantial disease caused by CC5-MRSA in human populations, the CC5 background has adapted to bird species and become a source of infection among broiler poultry (Lowder et al., 2009). The CC5 background also has been isolated from pigs (Frana et al., 2013), and with some traits that are present in other pig-adapted MRSA backgrounds (Hau et al., 2015).

The phylogenetic relationships between the numerous CC5-MRSA clones have not been determined, nor have the evolutionary events that led to their emergence. CC8 is the rival of CC5 for prominence among MRSA in the Western Hemisphere, and CC8-MRSA clone relationships and evolution have been studied extensively through genome sequencing (Uhlemann et al., 2014; Alam et al., 2015; Planet et al., 2015; Glaser et al., 2016; Challagundla et al., 2018). Relatively small samples of CC5-MRSA have been similarly studied (Kos et al., 2012; Aanensen et al., 2016; Arias et al., 2017). Here, we provide the first phylogenomic study of CC5-MRSA that is focused on large samples from the Western Hemisphere. Our goals are to resolve the phylogeny, place and time of origin of major CC5-MRSA clones, and to trace the evolution of their key traits.

MATERIALS AND METHODS

Bacterial Samples and Genome Sequencing

The genome sequences of 598 strains of CC5 were studied here. Sequences were newly generated for 409 strains, and published sequences were downloaded from publicly available sources for 189 strains. The newly generated sequences were from three sequencing facilities: the Broad Institute ($n = 297$), the University of Mississippi Medical Center (UMMC, $n = 74$), and Universidad El Bosque (UEB, $n = 38$). Information on these sequences is provided in **Supplementary Table S1**. All new sequences were generated using Illumina instruments. The Broad Institute's sequencing procedures are described in **Supplementary Text S1**, whereas previously published procedures were used at UMMC (Challagundla et al., 2018) and UEB (Arias et al., 2017). Eleven countries in the Western Hemisphere were sampled: Argentina, $n = 30$; Brazil, $n = 23$; Canada, $n = 36$; Chile, $n = 23$; Colombia, $n = 56$; Ecuador, $n = 9$; Guatemala, $n = 21$; Mexico, $n = 40$; Peru, $n = 28$; United States, $n = 270$; and Venezuela, $n = 21$. The US samples consisted of separate large collections: California, $n = 40$; Massachusetts, $n = 40$; Mississippi, $n = 89$; New York, $n = 74$; plus reference strains of USA100 and USA800, vancomycin-resistant *S. aureus* (VRSA), and avian and porcine strains. Other strains included completely sequenced reference strains JH1, N315, ED98, CF-Marseille, and 32 strains from Europe. Dates of isolation ranged from 1964 to 2017. Most strains were from clinical specimens.

Genome Assembly, Pseudoread Generation, and Alignment

Some of the published sequences were available only as assembled genomes. For those published sequences available as reads, and for the new sequences generated by UMMC, assemblies were made after filtering reads for minimum quality (base quality $\geq$ Q13, number of ambiguities $\leq$ 2, read length $\geq$ 15) using CLC Genomics Workbench v7 (Qiagen, Aarhus, Denmark). For those new sequences generated by the Broad Institute, assembly was carried out using the ALLPATHS-LG pipeline[1], followed by automated assembly improvement using Pilon[2]. Quality measures of all assemblies are listed in **Supplementary Table S1**.

In order to align this heterogeneous sample of sequences, we first used wgsim v0.3.2 (Li et al., 2009) with each assembly to generate 500,000 paired-end pseudoreads of 100 bp, with a 200 bp distance between the ends. These pseudoreads were mapped to the complete reference sequence of CC5 strain JH1 (Mwangi et al., 2007) using bwa v0.7.12 (Li and Durbin, 2010), coordinate-sorted with Picard v1.141 (DePristo et al., 2011), and realigned around indels with GATK v2.8-1 (DePristo et al., 2011).

Variant Calling and Functional Effects

Single-nucleotide polymorphisms (SNPs) and short insertion–deletion polymorphisms (indels) were called with the UnifiedGenotyper walker of GATK. All self-similar sequences identified by pairwise megablast of the CC5 reference sequence against itself, plus five regions of mobile genetic elements, were excluded from variant calling. Mapping quality was >50 for all variants. The functional effects of the variants were determined using SnpEff v4 (Cingolani et al., 2012) with the annotation of the CC5 reference sequence.

Strain Typing and Prediction of Antibiotic Resistance

Multilocus sequence typing (MLST) was done by scanning the strains' genome assemblies against the *S. aureus* MLST database[3], using the mlst v2.10 program[4] with default settings. Typing of the SCC*mec* element was done by mapping the strains' pseudoreads to a custom-clustered database of *ccr* and *mec* gene complexes, plus SCC*mec* IV subtype-specific sequences (Kaya et al., 2018), using the SRST2 v0.2.0 program (Inouye et al., 2014) with the min_coverage 60 option. The validity of this approach was investigated using pseudoreads generated from reference sequences of SCC*mec*[5] and by manually checking non-typeable elements with the SCC*mec*Finder web-based tool (Kaya et al., 2018). The detection of antibiotic resistance genes and mutations and the corresponding predicted antibiotic resistance was done with Mykrobe predictor v0.1.3, which has shown high (>99%) sensitivity and specificity for predicting resistance of *S. aureus* to a panel of 12 antibiotics (Bradley et al., 2015).

Detection of Selected Virulence Factors and Mobile Genetic Elements

To detect virulence factors, the full nucleotide dataset for *S. aureus* available at the Virulence Factors Database (Chen et al., 2016) was downloaded and then clustered into 265 groups with CD-HIT, using the tools available with SRST2 with default settings. The strains' pseudoreads were mapped to this clustered database using SRST2 with the min_coverage 60 option. A similar approach was used to detect mobile genetic elements, specifically phages and integrative conjugative elements (ICEs) that integrate into the *S. aureus* chromosome. Phages were detected based on their integrase genes, using the 12 integrase groups described by Goerke et al. (2009) plus the integrase for φSPβ (Holden et al., 2010). ICEs were detected based on full-length sequences, using the seven subfamilies of ICE*6013* and two subfamilies of Tn*916* (Tn*916* and Tn*5801*) described by Sansevere and Robinson (2017) and Sansevere et al. (2017). The validity of this approach was investigated using pseudoreads generated from complete sequences of reference strains of CC1, CC5, CC8, CC30, CC151, and reference sequences of the phages and ICEs.

---

[1]http://www.broadinstitute.org/software/allpaths-lg/blog

[2]http://www.broadinstitute.org/software/pilon/

[3]https://pubmlst.org/saureus/

[4]https://github.com/tseemann/mlst

[5]http://www.sccmec.org

Phylogenetic Inference and Recombination Detection

A maximum-likelihood (ML) phylogeny was constructed using PhyML v3.0 (Guindon et al., 2009) with the HKY+G model and five random and one BioNJ starting trees. For this analysis, an alignment representing biallelic SNPs and invariant core sequence present in all 598 genomes was used. Branch support was estimated with SH-like tests. The root of the tree was determined by first mapping reads from CC1, CC8, and ST72 outgroup sequences to the CC5 reference sequence, and then extracting the positions of the CC5 biallelic SNPs from the outgroups. These SNPs were then analyzed with distance-based (BioNJ) and maximum parsimony (MP) phylogenetic analysis. The outgroups consistently bisected a branch leading to a strain from Argentina, which was designated as the root. Recombination was detected using the PhyML tree with the method implemented in ClonalFrameML (CFML; Didelot and Wilson, 2015). Clade-specific recombination rates were examined on the recombination-corrected ML tree using R scripts.

Molecular Clock Analysis

The R package of Murray et al. (2016) was used to verify that no confounding existed between the path lengths on the recombination-corrected ML tree and the temporal distances of the genomes, and to verify a temporal phylogenetic signal in the positive relationship between the recombination-corrected ML tree root-to-tip distances and year of isolation of genomes. To estimate rates of evolution and dates of selected most recent common ancestors (MRCAs; tree nodes), the BEAST v1.7.5 program (Drummond et al., 2012) was used. We applied an HKY+G substitution model with empirical base frequencies, a Bayesian Skyline demographic model with default parameters, and both strict molecular clock and uncorrelated lognormal-relaxed molecular clock models using flat priors between 10^{-3} and 10^{-9} substitutions/site/year as informed by other *S. aureus* studies (Smyth et al., 2010). The basal strains from Argentina were constrained as outgroups, by forcing all other strains to be a monophyletic ingroup. Due to limitations in computational resources, the BEAST runs used three subsamples, each of 101 genomes. Each subsample consisted of 44 genomes chosen to represent early isolates and major nodes on the ML tree, plus 57 randomly selected genomes. Each subsample was run three times, and each run was 100,000,000 steps with sampling every 10,000 steps. Convergence and mixing of the MCMC chains, and the effective sample sizes of parameters (>2374 for the clock.rate parameter for every run of the selected strict clock model), was checked using Tracer v1.6. The first 10% of samples were removed as burn-in, the remaining samples were combined, and maximum clade credibility (MCC) trees were generated using LogCombiner v1.7.5 and TreeAnnotator v1.7.5. Since rate heterogeneity among branches was not extreme in any run of the relaxed clock model (i.e., the 95% credibility interval for the ucld.sd and coeeffvar parameters did not include 1.0), the strict clock model was selected.

Ancestral State Reconstruction

The evolution of selected traits was studied using ancestral state reconstruction under ML models. The ML analysis was done using the ace function of the ape R package (Paradis et al., 2004). Discrete, two-state characters were investigated including: geography (Central and North America vs. South America), predicted antibiotic resistance (resistant vs. susceptible), and the presence vs. absence of virulence factors, phages, ICEs, and predicted high-impact variants (i.e., frameshifts, start or stop codons gained or lost). Only those traits present in 5% ($n = 30$) to 95% ($n = 568$) of the genomes were investigated. The recombination-corrected ML tree was used as the estimate of the phylogeny. For each trait, both equal rates and all-rates-different models were fit to the data, and models were compared with a likelihood ratio test with one degree of freedom. In no case was the simpler, equal rates model rejected in favor of the more complicated, all-rates-different model.

Statistical Analysis and Tree Presentation

The basic stats package of R v3.4.2 (R Development Core Team, 2014) was used for two-sided tests of equal means (*t*-tests) and of no correlation between variables (Pearson's coefficient). Tajima's *D*, which examines the variance-normalized difference between two estimators of genetic diversity that are sensitive to different population dynamics, was calculated with the pegas (Paradis, 2010) R package. Tests of $D = 0$ were done assuming a beta distribution of the test statistic. Trees were visualized with iTOL (Letunic and Bork, 2016), and the ggtree (Yu et al., 2017) and phytools (Revell, 2012) R packages.

RESULTS AND DISCUSSION

Outline of the CC5 Phylogeny

The alignment of 598 genomes of CC5 consisted of 1,081,440 bp present in all genomes, which included 11,961 biallelic SNPs. ML phylogenetic analysis with these sites, followed by correction of tree branch lengths for recombination and outgroup-rooting, resulted in the phylogeny presented in **Figure 1**. Several clades were defined by consideration of their ST-SCC*mec* types, geographic distribution, and statistical support. These clades are described below.

CC5-Basal

CC5-Basal was defined as a large paraphyletic clade with a mixture of early-branching MSSA and MRSA predominantly of ST5-IV (SCC*mec* subtypes IVa, IVc, IVg, IVi) (**Figure 1**, light gray shading). This clade was well-dispersed across the Western and Eastern Hemispheres (**Figure 1**, ring 1). The most basal strain in our sample was an MSSA from Argentina, and the next most basal strains were a subclade of ST5-IVa (and one variant of ST5) from Argentina with the *pvl* and *sea* toxin genes that are characteristic of a recently emerged community-associated clone (Sola et al., 2008). Despite the exclusively South American source of these basal strains, ML ancestral state

FIGURE 1 | Phylogeny of CC5. The ML phylogeny has recombination-corrected branch lengths and is outgroup-rooted. The major clades of CC5-Basal, CC5-I, and CC5-II are indicated with shading. The statistical support for these clades is described in the text. The positions of reference strains are indicated with stars. The four rings in the direction of inner-outer, show continent (ring 1), country (ring 2), SCC*mec* type (ring 3), and multilocus ST (ring 4) for each genome. STs with less than three genomes were combined into "Other STs" for display purposes. Scale bar indicates number of substitutions per site.

reconstruction did not reliably identify the continent of origin of CC5 (**Supplementary Figure S1**). The Bayesian phylogenetic analysis constrained these basal strains as an outgroup, resulting in maximal posterior probability (PP = 1) for the node, but the ML phylogenetic analysis was unconstrained and still strongly supported the node (SH-like support = 0.984). The Bayesian analysis estimated the MRCA of this sample of CC5 as the late 1930s (**Table 1**).

CC5-Basal included the ST5-IVc USA800 clone from the United States (McDougal et al., 2003), and the ST5-I Geraldine clone from France (Dauwalder et al., 2008). Also included are VRSA strains HOU1444-VR from Brazil and VRS3a from the United States, which represented independent acquisitions of vancomycin resistance and confirmed previous analysis (Panesso et al., 2015). CC5-Basal further included ST5-IVa reference strain CF-Marseille from France (Rolain et al., 2009), ST5-II reference strain N315 from Japan (Kuroda et al., 2001), and rare strains of ST5-I and ST5-III from Europe. The Bayesian and ML analyses provided maximal support (PP = 1, SH-like support = 1) for a previously described avian subclade, which

included reference strain ED98 (Lowder et al., 2009), and for a separate subclade from porcine sources (Hau et al., 2017). The ML analysis presented a CC5-Basal subclade that included the porcine subclade (**Figure 1**, light gray shading), as a sister to the CC5-II clade described below. However, neither the Bayesian nor ML analyses provided statistical support for this arrangement, so this subclade should be considered to be part of the broader CC5-Basal clade.

CC5-I

CC5-I was defined as a large monophyletic clade with SCC*mec* I (**Figure 1**, light red shading). This clade represented one of three separate acquisitions of SCC*mec* I within CC5—the others occurred within the CC5-Basal clade. Most early-branching strains of CC5-I were from Europe, the rest were from South America. The ST228-I South German clone (Ghebremedhin et al., 2005) and related STs ST111 and ST1481 (Budimir et al., 2010) represented these early-branching European strains of CC5-I. The likelihood of a South American origin among CC5-I strains in the Western Hemisphere was

TABLE 1 | Rates of evolution and dates of CC5 clades.

Clade or subclade	Date of MRCA (95% credibility interval)[a]	Ratio of recombination to mutation events, ρ/θ (95% credibility interval)[b]
All CC5	1938 (1930, 1945) 1937 (1928, 1944) 1938 (1930, 1945)	0.0046 (0.0035, 0.0059)
CC5-Basal	Same as all CC5	0.0016 (0.00063, 0.0029)
CC5-I	1973 (1969, 1977) 1973 (1968, 1977) 1973 (1969, 1977)	0.0046 (0.0019, 0.0086)
…CC5-I expansion	1994 (1992, 1996) 1995 (1993, 1997) 1994 (1992, 1996)	0.0050 (0.0018, 0.0096)
CC5-II	1962 (1957, 1966) 1961 (1955, 1966) 1962 (1956, 1966)	0.0069 (0.0050, 0.0090)
…CC5-II A	Same as CC5-II	0.014 (0.0095, 0.019)
……Mexican subclade	1997 (1994, 2000) 1997 (1994, 1999) 1997 (1994, 2000)	0.019 (0.0089, 0.032)
…CC5-II B	1975 (1970, 1979) 1975 (1970, 1979) 1975 (1970, 1979)	0.0031 (0.0017, 0.0050)
……CC5-II B expansion	1982 (1978, 1986) 1982 (1978, 1986) 1982 (1979, 1986)	0.0029 (0.0014, 0.0049)

[a]For each clade or subclade, the three results for Date of Most Recent Common Ancestor (MRCA) refer to the three subsamples (see the section "Materials and Methods" for subsampling procedures). The three subsamples had estimated mutation rates (95% credibility intervals) in units of substitutions/site/year of 1.55×10^{-6} (1.39×10^{-6}, 1.71×10^{-6}), 1.52×10^{-6} (1.36×10^{-6}, 1.68×10^{-6}), 1.55×10^{-6} (1.39×10^{-6}, 1.70×10^{-6}), respectively. [b]The ratio of recombination to mutation events was calculated from the complete clade or subclade, without subsampling, using CFML.

>98% (**Supplementary Figure S1**). Bayesian and ML analyses provided maximal support for CC5-I, and its MRCA was estimated as the early 1970s (**Table 1**).

The terminal subclades of CC5-I represented an expansion of the ST5-I Chilean/Cordobes clone (Sola et al., 2002) throughout South America. The expansion node had maximal support in Bayesian and ML analyses, and its MRCA was estimated as the mid-1990s (**Table 1**). The initial branching events of the expansion were unresolved, which was consistent with a rapid expansion where the rate of geographic spread had outpaced the rate of mutation. Moreover, Tajima's D for strains of the expansion was -2.75 ($P = 3.46 \times 10^{-5}$), which was consistent with a recent expansion because it indicated an excess of rare alleles that would be purged from an older equilibrium population. Subsequent subclades were strongly structured geographically by country within South America, and included strains from Brazil, Chile and nested Argentina strains, Peru, Argentina, Colombia, and nested Venezuela strains (**Figure 1**, ring 2).

CC5-II

CC5-II was defined as a very large monophyletic clade with SCC*mec* II (**Figure 1**, light blue shading). This clade represented

one of two separate acquisitions of SCC*mec* II within CC5—the other occurred within the CC5-Basal clade on the branch leading to reference strain N315. CC5-II was subdivided into a paraphyletic early-branching subclade, CC5-II A, and a monophyletic terminal subclade, CC5-II B. In the Western Hemisphere, CC5-II strains were mostly from Central and North America, but some strains were from Brazil, Ecuador, and Venezuela. The likelihood of a Central or North American origin among CC5-II strains in the Western Hemisphere was >99% (**Supplementary Figure S1**). This clade had maximal support in Bayesian and ML analyses, and subclade CC5-II B had near maximal support (PP = 1, SH-like support = 0.997). The MRCAs of CC5-II and CC5-II B, respectively, were estimated as the early 1960s and mid-1970s (**Table 1**).

Of the four sampled reference strains of the ST5-II USA100 clone from the United States, one was within CC5-II A and three were within CC5-II B (**Figure 1**), which indicated that this clone was polyphyletic. Furthermore, the separation of ST5-II strains from New York in the United States and from Japan into distinct clades (CC5-II B and CC5-Basal, respectively), indicated that the New York/Japan clone was polyphyletic; in other words, there are separate New York and Japan clones, and not a single New York/Japan clone. Our study focused sampling on the Western Hemisphere, but it will be interesting to see if future studies place Asian SCC*mec* II-positive CC5 strains with strain N315 in CC5-Basal or with the diverse strains in CC5-II. Lastly, of the 11 sampled ST5-II VRSA strains from the United States, five were within CC5-II A and six were within CC5-II B (**Figure 1**, brown stars). Our analysis supported a previous analysis that indicated independent acquisitions of vancomycin resistance for most strains (Kos et al., 2012), but our analysis also highlighted the close relationships of strain pairs VRS1/VRS6, VRS5/VRS7, VRS9/VRS10, and VRS11a/VRS11b. Importantly, no evidence of further dissemination of these VRSA strains was obtained, as ML models of ancestral state reconstruction did not identify strains other than the known VRSAs to descend from vancomycin-resistant MRCAs (nodes). In the absence of vancomycin, the slight fitness burden reported for carriage of the *vanA* operon (Foucault et al., 2009) may be sufficient to impede the dissemination of these strains.

A major feature of CC5-II A was a divergent subclade that included all of the sampled strains from Mexico (**Figure 1**). Early-branching strains of this subclade were from the United States. This subclade had near maximal support (PP = 1, SH-like support = 0.999) in Bayesian and ML analyses, and its MRCA was estimated as the late 1990s (**Table 1**). This time period coincided with the replacement of CC30-MRSA in Mexico by CC5-MRSA, as previously described (Aires De Sousa et al., 2001; Velazquez-Meza et al., 2004; Echaniz-Aviles et al., 2006). This subclade was further subdivided into ST5-II and ST1011-II subclades that showed geographic structure by region within Mexico: ST5-II was from Central and Western Mexico, and ST1011-II was from Central and Northern Mexico.

The ST225-II Rhine-Hesse clone from Central Europe (Schulte et al., 2013) was an early offshoot of CC5-II B. The ST5-II reference strain JH1 from the United States (Mwangi et al., 2007) also nested within CC5-II B. The terminal subclades of CC5-II B

represented an expansion of ST5-II, and subsequent proliferation of unnamed clones such as ST105-II, ST125-II, ST231-II, and ST496-II (**Figure 1**, ring 3), in Central and North America. The expansion node had maximal support in Bayesian and ML analyses, and dated to the early 1980s (**Table 1**). Several of the initial branching events of the expansion were unresolved, which was consistent with a rapid expansion. Also, Tajima's D was -2.81 ($P = 5.91 \times 10^{-6}$), which was consistent with a recent expansion. CC5-II B showed some geographic structure at the country level, such as a subclade with 18 of 21 strains from Guatemala, two subclades with 10–11 strains each from Canada, and a subclade with 18 strains from California in the United States (**Figure 1**, ring 2).

Relatively Low Recombination Rates in CC5

Our estimate of the ratio of recombination to mutation events in this sample of CC5 (0.0046) (**Table 1**) was several orders of magnitude lower than previously estimated for CC5 (1.08) (Murray et al., 2017). However, the previous estimate detected a large proportion of recombination events in poultry strains and our sample included only six poultry strains for which no recombination events were detected. Sampling differences and the use of different recombination detection methods might explain the discrepancy between our recombination estimate and that of Murray et al. (2017). Our recombination estimate for CC5 was also lower than previously reported for the non-CC5 clones of ST239-III (Castillo-Ramirez et al., 2012) and ST8-IV USA300 (Challagundla et al., 2018); ranges of 0.05–0.29 and 0.12–0.29, respectively. The relatively low recombination rate estimated here using CFML was confirmed with an independent analysis using Gubbins (Croucher et al., 2015): 58 recombination events were detected by CFML and 29 were detected by Gubbins. In our sample, CC5-II A had a significantly higher ratio of recombination to mutation events compared to other CC5 clades (**Table 1**). This result was driven by a relatively large number of recombination events in the Mexican subclade. Thus, our analysis did not support the hypothesis that CC5 was generally more promiscuous than other *S. aureus* genetic backgrounds. Other explanations for why VRSA strains have repeatedly appeared in CC5 should be sought, such as whether CC5 has a unique gene regulatory environment that favors *vanA* expression or a unique ecological niche that favors interactions with *Enterococcus* spp.

Evidence for an Antibiotic Resistance-Toxicity Tradeoff in CC5

The full range of detected mutations and genes associated with antibiotic resistance, and selected virulence factors and mobile genetic elements, is provided in **Supplementary Table S1**. Our discussion is focused on those traits that are variably distributed among CC5 clades in the Western Hemisphere. Besides the high prevalence of resistance to β-lactams in all clades, CC5-I and CC5-II had a high prevalence of resistance to fluoroquinolones, macrolides, and lincosamides, and CC5-I also had a high prevalence of resistance to aminoglycosides (**Table 2**). The most common resistance mechanism to fluoroquinolones was

the "double-serine mutations," Ser84Leu in *gyrA* and Ser80Phe in *grlA*, that have been noted to occur in other successful MRSA clones (Fuzi et al., 2017). These double-serine mutations accounted for fluoroquinolone resistance in 141 of 147 (96%) resistant strains of CC5-I, and 107 of 239 (45%) resistant strains of CC5-II. Of note, CC5-II B had a lower prevalence of fluoroquinolone resistance compared to CC5-II A (58% vs. 98%, respectively) (**Table 2**), and the main mechanism of resistance in CC5-II B was the single mutation in *gyrA* that accounted for 117 of 141 (83%) resistant strains of CC5-II B. The most common resistance mechanism to macrolides and lincosamides was the *ermA* gene: alone accounting for 139 of 144 (97%) resistant strains of CC5-I, and 268 of 342 (78%) resistant strains of CC5-II. The *ermA* gene is known to be carried on Tn*554* on the SCC*mec* II element (Kuroda et al., 2001), and it occurred on Tn*554* elsewhere in the chromosome of the examined strains with SCC*mec* I elements. Aminoglycoside resistance, which occurred mostly in CC5-I, was exclusively attributed to the *aacA*–*aphD* bifunctional gene in our sample, and this gene occurred on composite transposon Tn*4001* in examined strains. Mobile genetic elements also carried some rare antibiotic resistance genes. For example, ICE element (also known as conjugative transposon) Tn*916* accounted for all eight strains with *tetM*-mediated tetracycline resistance. The other family of ICE elements known in *S. aureus*, ICE*6013*, was not observed to carry any antibiotic resistance genes and was common among the avian subclade of CC5-Basal (**Table 2**).

Common virulence factors of CC5 included the *sak*, *scn*, and *chp* genes of the immune evasion gene cluster (IEC) present on phage φSa3, and *seg*, *sei*, *sem*, *sen*, *seo*, and ψ*ent* of the enterotoxin gene cluster (EGC) as well as *lukDE* present on genomic island νSaβ (**Table 2**). CC5-Basal appeared to have a more diverse array of toxins, and the *pvl* (*lukFS*), *sec*, *sel*, and *etb* toxins occurred solely in this clade. CC5-I was unique in having a low prevalence of the *fnbB* adhesin gene; it was present in only 3% of CC5-I strains, and in >96% of strains of other clades (**Table 2**). On the other hand, the plasmid-borne *sed*, *sej*, *ser* toxins were most common in CC5-II B strains. Phages also showed some differences in CC5 clade distribution: phages φSa2 and φSa7 were most common in CC5-I strains, and phage φSa1 was most common in CC5-II B strains (**Table 2**).

We tested the hypothesis that antibiotic resistances and toxins were randomly distributed throughout CC5 and across time by comparing the average number of antibiotic resistances and toxins per genome in the various clades and by examining these traits with distance from the root of the CC5 tree. CC5-Basal had significantly fewer antibiotic resistances and more toxins than the other clades (**Table 3**). CC5-I had the opposite pattern, significantly more antibiotic resistances and fewer toxins than the other clades. CC5-II also had significantly more antibiotic resistances, but it had more toxins due to greater toxin acquisition in CC5-II B (e.g., *sed*, *sej*, *ser*) that offset the loss in CC5-II A (**Table 3**). Overall, the number of antibiotic resistances per genome increased with distance from root, and the number of toxins per genome decreased with distance from root (**Table 3**). A similar trend of increased number of antibiotic resistances with distance from root has been reported previously for CC8

TABLE 2 | Prevalence of selected traits in CC5 clades.

	Trait	Clade or subclade[a]			
		CC5-Basal (*n* = 107)	CC5-I (*n* = 148)	CC5-II A (*n* = 100)	CC5-II B (*n* = 243)
Antibiotic resistances	CIP	33 (30.84)	147 (99.32)	98 (98)	141 (58.02)
	ERY	28 (26.17)	144 (97.3)	99 (99)	243 (100)
	CLIN	15 (14.02)	144 (97.3)	99 (99)	243 (100)
	FUS	1 (0.93)	1 (0.68)	0 (0)	3 (1.23)
	GENT	16 (14.95)	143 (96.62)	19 (19)	13 (5.35)
	OXA	63 (58.88)	148 (100)	99 (99)	239 (98.35)
	MUP	1 (0.93)	1 (0.68)	6 (6)	10 (4.12)
	PEN	89 (83.18)	148 (100)	100 (100)	241 (99.18)
	RIF	8 (7.48)	11 (7.43)	7 (7)	10 (4.12)
	TET	20 (18.69)	1 (0.68)	1 (1)	4 (1.65)
	TMP	10 (9.35)	1 (0.68)	16 (16)	12 (4.94)
	VANC	2 (1.87)	0 (0)	5 (5)	6 (2.47)
Virulence factors	*etb*	9 (8.41)	0 (0)	0 (0)	0 (0)
	lukDE	104 (97.2)	145 (97.97)	89 (89)	229 (94.24)
	lukFS	6 (5.61)	0 (0)	0 (0)	0 (0)
	sea	45 (42.06)	16 (10.81)	38 (38)	3 (1.23)
	seb	7 (6.54)	0 (0)	0 (0)	6 (2.47)
	sec	4 (3.74)	0 (0)	0 (0)	0 (0)
	sed	20 (18.69)	0 (0)	24 (24)	163 (67.08)
	seg	106 (99.07)	142 (95.95)	90 (90)	226 (93)
	sei	106 (99.07)	144 (97.3)	87 (87)	226 (93)
	sej	23 (21.5)	0 (0)	23 (23)	163 (67.08)
	sek	2 (1.87)	1 (0.68)	0 (0)	1 (0.41)
	sel	4 (3.74)	0 (0)	0 (0)	0 (0)
	sem	107 (100)	144 (97.3)	91 (91)	231 (95.06)
	sen	106 (99.07)	139 (93.92)	89 (89)	225 (92.59)
	seo	107 (100)	146 (98.65)	91 (91)	231 (95.06)
	sep	58 (54.21)	7 (4.73)	63 (63)	12 (4.94)
	seq	2 (1.87)	0 (0)	0 (0)	1 (0.41)
	ser	22 (20.56)	1 (0.68)	23 (23)	163 (67.08)
	sev	107 (100)	144 (97.3)	87 (87)	228 (93.83)
	tst1	12 (11.21)	1 (0.68)	1 (1)	1 (0.41)
	ψ*ent2*	106 (99.07)	143 (96.62)	90 (90)	228 (93.83)
	fnbB	105 (98.13)	5 (3.38)	97 (97)	238 (97.94)
	sdrE	99 (92.52)	143 (96.62)	89 (89)	229 (94.24)
	chp	70 (65.42)	118 (79.73)	85 (85)	223 (91.77)
	scn	85 (79.44)	133 (89.86)	91 (91)	226 (93)
	sak	86 (80.37)	133 (89.86)	92 (92)	230 (94.65)
Phages	ϕSa12*int*	1 (0.93)	0 (0)	0 (0)	2 (0.82)
	ϕSa1*int*	39 (36.45)	3 (2.03)	41 (41)	194 (79.84)
	ϕSa2*int*	31 (28.97)	110 (74.32)	61 (61)	132 (54.32)
	ϕSa3*int*	92 (85.98)	134 (90.54)	90 (90)	231 (95.06)
	ϕSa4*int*	10 (9.35)	2 (1.35)	16 (16)	13 (5.35)
	ϕSa5*int*	6 (5.61)	2 (1.35)	6 (6)	66 (27.16)
	ϕSa6*int*	14 (13.08)	6 (4.05)	40 (40)	20 (8.23)
	ϕSa7*int*	18 (16.82)	147 (99.32)	65 (65)	75 (30.86)
	ϕSa9*int*	0 (0)	0 (0)	1 (1)	0 (0)
	ϕSPβ*int*	0 (0)	0 (0)	0 (0)	3 (1.23)
ICEs	ICE*6013*_sf1	23 (21.49)	1 (0.68)	6 (6)	6 (2.47)
	ICE*6013*_sf5	8 (7.48)	0 (0)	0 (0)	0 (0)
	ICE*6013*_sf6	6 (5.61)	0 (0)	0 (0)	0 (0)
	Tn*916*	6 (5.61)	0 (0)	0 (0)	2 (0.82)

[a]*Numbers in cells refer to the number (percentage) of genomes in the indicated clade or subclade that have the indicated trait.*

TABLE 3 | Comparisons of antibiotic resistance and toxin traits in CC5 clades.

Clade or subclade	Number of genomes	Average number per genome (SD)[a]		Correlation between number per genome and root-to-tip distance[b]	
		Antibiotic resistances[c]	Toxins[d]	Antibiotic resistances[c]	Toxins[d]
All CC5	598	4.8 (1.6)	9.2 (2.1)	0.394****	−0.326****
CC5-Basal	107	2.7 (2.0)****	9.9 (2.1)****	0.029	−0.295**
CC5-I	148	6.0 (0.6)****	7.9 (1.0)****	−0.178*	0.126
CC5-II	343	5.0 (0.9)**	9.4 (2.3)***	−0.017	−0.194***
. . .CC5-II A	100	5.5 (1.0)****	8.9 (2.4)	−0.224*	−0.037
. . .CC5-II B	243	4.8 (0.8)	9.7 (2.2)****	−0.089	−0.284****

[a]Standard deviation (SD), test is for equal means of clade or subclade vs. all others. [b]Test is for zero correlation, root-to-tip distance is from CFML tree. [c]From a possible total of 12 predicted resistances (CIP, CLIN, ERY, FUS, GENT, OXA, MUP, PEN, RIF, TET, TMP, VANC). [d]From a possible total of 20 toxins (etb, lukDE, lukFS, sea, seb, sec, sed, seg, sei, sej, sek, sel, sem, sen, seo, sep, seq, ser, sev, tst1). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

(Strommenger et al., 2014). While these results highlight the importance of antibiotic resistance to the evolution of CC5, they are subject to a statistical caveat. Not all of the data points are independent because mobile genetic elements can carry multiple linked antibiotic resistance or toxin genes and affect multiple traits in a single evolutionary event; for example, *ermA* mediates both macrolide and lincosamide resistance, and the EGC encodes multiple enterotoxins.

## Convergent Genomic Changes Associated With Expansions of CC5

The evolution of selected traits was studied in more detail under an ML model of ancestral state reconstruction. A major finding of this analysis was that resistance to fluoroquinolones, macrolides, and lincosamides was gained, and the *sep* toxin gene was lost, independently, along the phylogenetic backbone leading to the expansion nodes of CC5-I and CC5-II B (**Figure 2**, pie-charts). The likelihood of the nodes that precede CC5-I and CC5-II B having these resistances was <5–6% for fluoroquinolones and <1% for macrolides and lincosamides, and >98% for the presence of *sep*. By the time of the expansion nodes of CC5-I and CC5-II B, the likelihood of having all three resistances was >99%, and <1% for the presence of *sep* (**Figure 2**, pie-charts).

The independent evolution of the same traits from different starting points on the CC5 phylogeny (i.e., convergence), and the close timing of their evolution with independent expansions, suggests a causal relationship. Use of β-lactams, macrolides, and fluoroquinolones rank highly among antibiotics in the United States, Canada, and Brazil (Center for Disease Dynamics et al., 2015), and likely other countries in South America. Thus, the gain of these resistances may reflect a common selective pressure in the Western Hemisphere. As can be seen in **Figure 2** (bar charts), some of the strains that appeared after the expansion of CC5-II B have subsequently lost fluoroquinolone resistance. As indicated above, many of the CC5-II B strains that retained fluoroquinolone resistance have retained the *gyrA* resistance allele but have lost the *grlA* resistance allele. Fluoroquinolone resistance mutations and especially the *grlA* mutation, have been shown to negatively impact the fitness of *S. aureus* strains in the absence of antibiotic (Horváth et al., 2012; Knight et al., 2012).

However, these fitness effects may be mitigated in the presence of sub-inhibitory levels of the antibiotic (Gustave et al., 2018). These observations suggest that fluoroquinolone resistance may have an important role in the initial phase of epidemic spread CC5-MRSA clones, but less of a role in the subsequent phase of endemic residence in hospitals. Alternatively, there may have been efforts to reduce fluoroquinolone use in recent years, which in turn may have alleviated the selective pressure for clones to remain resistant. The CC5-II B expansion began approximately one decade earlier than the CC5-I expansion (**Table 1**), so it will be interesting to see if CC5-I begins to lose fluoroquinolone resistance and the *grlA* resistance allele over time as has happened in CC5-II B.

The association of the CC5-I and CC5-II B expansions with independent losses of the *sep* toxin gene was unexpected. This gene is present on the IEC of phage φSa3. Its loss preceding the CC5-I and CC5-II B expansions does not represent loss of the phage, since the majority of strains retained the phage and other virulence factors of the IEC such as *sak*, *scn*, and *chp* (**Table 2**). The completely sequenced reference strains N315 and JH1, respectively, provide full sequences of the IEC before and after loss of *sep*. A 1.75 kb region of the IEC that includes the *sep* gene was noted to be variably present in *S. aureus* by van Wamel et al. (2006). To our knowledge, the mechanism of *sep* excision from the IEC is unknown, and the function of *sep* is unstudied beyond its characterization as a superantigen with emetic activity (Omoe et al., 2005, 2013). In one study, the presence of *sep* was identified as the only *S. aureus* virulence factor among 30 tested, to associate with bacteremia in hospitalized MRSA carriers (Calderwood et al., 2014). While *sep* and other genome variations might influence the risk of bacteremia, those types of invasive infections are dead-ends for MRSA transmission.

The overall trend of decreased number of toxins per genome with distance from root, and parallel losses of *sep* in particular, suggests that the CC5 expansions occurred with strains that were less virulent than their precursors. Additional observations that support this notion come from an analysis of 35 high-frequency, high-impact mutations (listed in **Supplementary Table S1**). A total of 14 of these mutations occurred in CC5-I by the time of the expansion and three more occurred after
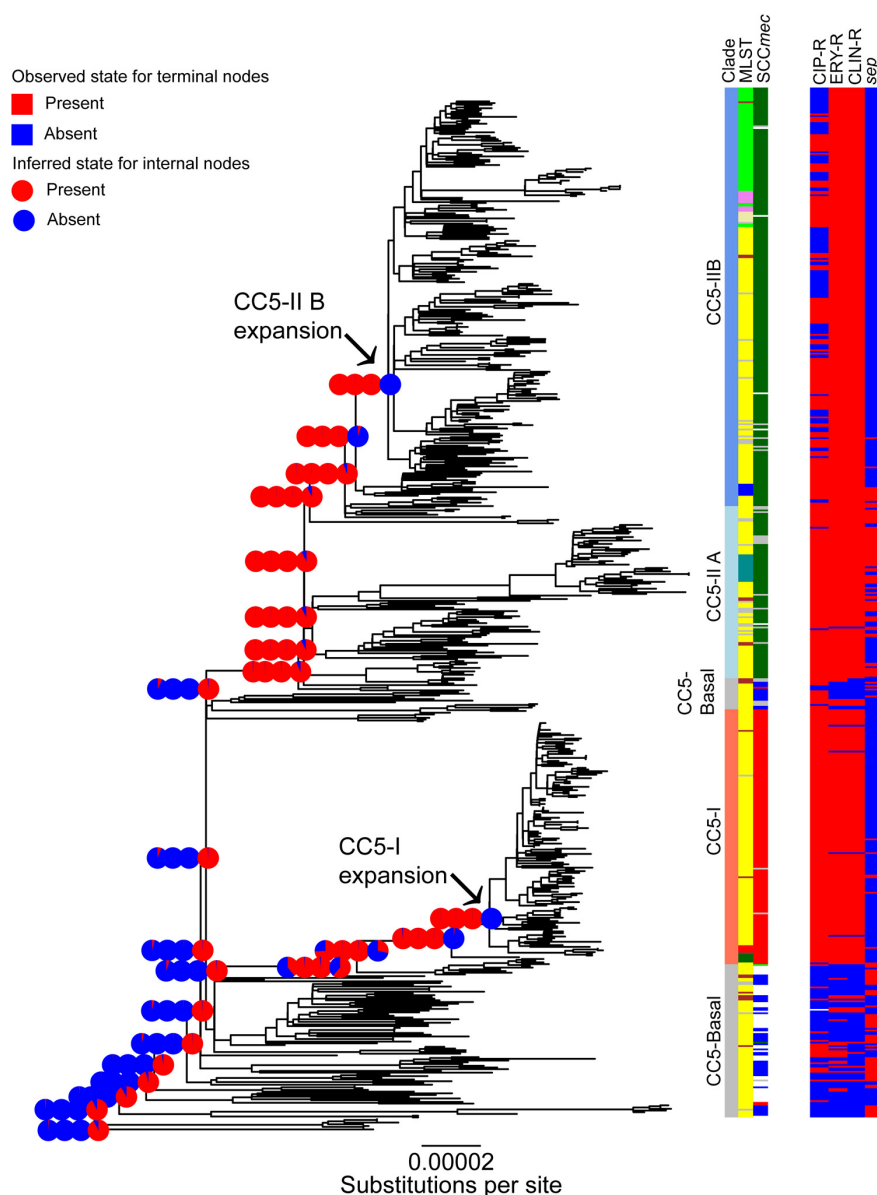
FIGURE 2 | Convergent evolution within CC5. The ML phylogeny of **Figure 1** is shown. The major clades, multilocus ST, and SCC*mec* type are shown for reference purposes and colored as in **Figure 1**. The CC5-I and CC5-II B expansion nodes are indicated with arrows. ML ancestral state reconstruction is shown along the phylogenetic backbone leading to the two expansion nodes; the four pie-charts at each of the backbone nodes indicate the likelihood of (in left-to-right order), fluoroquinolone, macrolide, and lincosamide resistance (red) or susceptibility (blue), and the presence (red) or absence (blue) of the *sep* toxin gene. Table adjacent to the tree indicates these traits in each genome. Scale bar indicates number of substitutions per site.

the expansion. One of these was a frameshift mutation in the *sasG* adhesin gene, which also occurred in CC5-I strains that lacked the *fnbB* gene. Loss of *sasG* and *fnbB* function would be expected to result in reduced virulence especially in biofilm-associated infections (Vergara-Irigaray et al., 2009; Geoghegan et al., 2010). In CC5-II B, only two high-frequency, high-impact mutations occurred by the time of the expansion, but four occurred afterward. One of these mutations resulted in the loss of the stop codon of the *srtA* sortase gene, which is a virulence factor that attaches surface proteins with the

LPXTG motif to *S. aureus*' peptidoglycan cell wall. Loss of sortases function would be expected to result in reduced virulence (Mazmanian et al., 2000). One interesting gain-of-function change in CC5-II B was a frameshift mutation that restored the start codon of the toxin component of the *axe1/txe1* toxin/antitoxin system. Expression of this system in CC8 strain Newman is increased after exposure to subinhibitory concentrations of erythromycin and tetracycline (Donegan and Cheung, 2009), but it has not been studied in CC5 to our knowledge.

CONCLUDING REMARKS

CC5-MRSA in the Western Hemisphere are highly diverse, even those strains that share the same ST-SCC*mec* type and circulate in the same country. More precise definitions for commonly sampled clones such as the USA100 and New York/Japan clones, and likely other clones such as USA800, are needed because different strains with those labels may have shared a MRCA > 50 years ago and may have divergent genomes. Our study provides the first systematic effort at organizing this diversity from a phylogenomic perspective and it provides a robust landmark for future genome studies of CC5-MRSA strains. The geographic structure of CC5 that is evident at the continent, country, and even region levels in some cases, suggests that a more precise delineation of the patterns of geographic spread of CC5-MRSA clones may be possible. MRSA clones of clinical relevance branch at nearly all time depths in the CC5 phylogeny; from the earliest-branching ST5-IV clones within the CC5-Basal clade that are an emerging problem in hospitals and communities in South America, to the latest-branching ST5-II clones within the CC5-II B clade that are a continuing problem in hospitals in Central and North America. Our analysis shows relatively low rates of recombination in CC5, which indicates that the propensity of CC5 to acquire *vanA*-mediated vancomycin resistance is not likely due to enhanced promiscuity and prompts study of other potential mechanisms. In tracing the evolution of selected traits of CC5, we discovered that some of the genomic changes that occurred prior to expansions of the CC5-I and CC5-II B clades represented instances of convergent evolution. While the convergent acquisition of resistance to widely prescribed antibiotics is a recurring theme in the history of successful MRSA clones, the convergent loss of the *sep* toxin and the trend of decreasing number of toxins with distance from the CC5 tree root and loss of other virulence factors is a unique finding. Taken together, our results suggest that more antibiotic-resistant and less virulent CC5-MRSA clones may be better able to spread geographically.

DATA AVAILABILITY

Sequence reads are available from NCBI Bioproject PRJNA224189, PRJNA454482, and PRJNA291213. Accession numbers for sequences are provided in **Supplementary Table S1**.

DATA DEPOSITION

NCBI Bioproject (http://www.ncbi.nlm.nih.gov/bioproject) PRJNA224189, PRJNA454482, and PRJNA291213.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01901/full#supplementary-material

FIGURE S1 | Geographic structure within CC5.

TABLE S1 | Characteristics of CC5 genome sequences.

TEXT S1 | Genome sequencing procedures for the Broad Institute.

REFERENCES

Aanensen, D. M., Feil, E. J., Holden, M. T., Dordel, J., Yeats, C. A., Fedosejev, A., et al. (2016). Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *Staphylococcus aureus* in Europe. *mBio* 7:e444-16. doi: 10.1128/mBio.00444-16

Aires De Sousa, M., Miragaia, M., Sanches, I. S., Avila, S., Adamson, I., Casagrande, S. T., et al. (2001). Three-year assessment of methicillin-resistant *Staphylococcus aureus* clones in Latin America from 1996 to 1998. *J. Clin. Microbiol.* 39, 2197–2205. doi: 10.1128/JCM.39.6.2197-2205.2001

Alam, M. T., Read, T. D., Petit, R. A., Boyle-Vavra, S., Miller, L. G., Eells, S. J., et al. (2015). Transmission and microevolution of USA300 MRSA in U.S. Households: evidence from whole-genome sequencing. *mBio* 6:e00054-15. doi: 10.1128/mbio.00054-15

Arias, C. A., Reyes, J., Carvajal, L. P., Rincon, S., Diaz, L., Panesso, D., et al. (2017). A prospective cohort multicenter study of molecular epidemiology and phylogenomics of *Staphylococcus aureus* bacteremia in nine Latin American Countries. *Antimicrob. Agents Chemother.* 61, e816–e817. doi: 10.1128/AAC.00816-17

Boucher, H., Miller, L. G., and Razonable, R. R. (2010). Serious infections caused by methicillin-resistant *Staphylococcus aureus*. *Clin. Infect. Dis.* 51(Suppl. 2), S183–S197. doi: 10.1086/653519

Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., et al. (2015). Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat. Commun.* 6:10063. doi: 10.1038/ncomms10063

Budimir, A., Deurenberg, R. H., Bosnjak, Z., Stobberingh, E. E., Cetkovic, H., and Kalenic, S. (2010). A variant of the Southern German clone of methicillin-resistant *Staphylococcus aureus* is predominant in Croatia. *Clin. Microbiol. Infect.* 16, 1077–1083. doi: 10.1111/j.1469-0691.2009.03042.x

Calderwood, M. S., Desjardins, C. A., Sakoulas, G., Nicol, R., Dubois, A., Delaney, M. L., et al. (2014). Staphylococcal enterotoxin P predicts bacteremia in hospitalized patients colonized with methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* 209, 571–577. doi: 10.1093/infdis/jit501

Castillo-Ramirez, S., Corander, J., Marttinen, P., Aldeljawi, M., Hanage, W. P., Westh, H., et al. (2012). Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol.* 13:R126. doi: 10.1186/gb-2012-13-12-r126

Center for Disease Dynamics, Economics and Policy (2015). *Antibiotic use in 2015*. Available at: https://resistancemap.cddep.org/AntibioticUse.php [accessed May 1, 2018].

Centers for Disease Control and Prevention (2013). *Antibiotic Resistance Threats in the United States, 2013*. Available at: https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf [accessed March 2, 2018].

Challagundla, L., Luo, X., Tickler, I. A., Didelot, X., Coleman, D. C., Shore, A. C., et al. (2018). Range expansion and the origin of USA300 North American epidemic methicillin-resistant *Staphylococcus aureus*. *mBio* 9:e02016-17. doi: 10.1128/mBio.02016-17

Chen, L., Zheng, D., Liu, B., Yang, J., and Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. *Nucleic Acids Res.* 44, D694–D697. doi: 10.1093/nar/gkv1239

Christianson, S., Golding, G. R., Campbell, J., Canadian Nosocomial Infection Surveillance Program, and Mulvey, M. R. (2007). Comparative genomics of Canadian epidemic lineages of methicillin-resistant *Staphylococcus aureus*. *J. Clin. Microbiol.* 45, 1904–1911. doi: 10.1128/JCM.02500-06

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6, 80–92. doi: 10.4161/fly.19695

Coombs, G. W., Van Gessel, H., Pearson, J. C., Godsell, M. R., O'Brien, F. G., and Christiansen, K. J. (2007). Controlling a multicenter outbreak involving the New York/Japan methicillin-resistant *Staphylococcus aureus* clone. *Infect. Control Hosp. Epidemiol.* 28, 845–852. doi: 10.1086/518726

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43:e15. doi: 10.1093/nar/gku1196

Dauwalder, O., Lina, G., Durand, G., Bes, M., Meugnier, H., Jarlier, V., et al. (2008). Epidemiology of invasive methicillin-resistant *Staphylococcus aureus*

clones collected in France in 2006 and 2007. *J. Clin. Microbiol.* 46, 3454–3458. doi: 10.1128/JCM.01050-08

David, M. Z., Daum, R. S., Bayer, A. S., Chambers, H. F., Fowler, V. G. Jr., Miller, L. G., et al. (2014). *Staphylococcus aureus* bacteremia at 5 US academic medical centers, 2008-2011: significant geographic variation in community-onset infections. *Clin. Infect. Dis.* 59, 798–807. doi: 10.1093/cid/ciu410

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806

Didelot, X., and Wilson, D. J. (2015). ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comp. Biol.* 11:e1004041. doi: 10.1371/journal.pcbi.1004041

Diekema, D. J., Richter, S. S., Heilmann, K. P., Dohrn, C. L., Riahi, F., Tendolkar, S., et al. (2014). Continued emergence of USA300 methicillin-resistant *Staphylococcus aureus* in the United States: results from a nationwide surveillance study. *Infect. Control Hosp. Epidemiol.* 35, 285–292. doi: 10.1086/675283

Donegan, N. P., and Cheung, A. L. (2009). Regulation of the mazEF toxin-antitoxin module in *Staphylococcus aureus* and its impact on sigB expression. *J. Bacteriol.* 191, 2795–2805. doi: 10.1128/JB.01713-08

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Echaniz-Aviles, G., Velazquez-Meza, M. E., Aires-de-Sousa, M., Morfin-Otero, R., Rodriguez-Noriega, E., Carnalla-Barajas, N., et al. (2006). Molecular characterisation of a dominant methicillin-resistant *Staphylococcus aureus* (MRSA) clone in a Mexican hospital (1999-2003). *Clin. Microbiol. Infect.* 12, 22–28. doi: 10.1111/j.1469-0691.2005.01283.x

Egea, A. L., Gagetti, P., Lamberghini, R., Faccone, D., Lucero, C., Vindel, A., et al. (2014). New patterns of methicillin-resistant *Staphylococcus aureus* (MRSA) clones, community-associated MRSA genotypes behave like healthcare-associated MRSA genotypes within hospitals, Argentina. *Int. J. Med. Microbiol.* 304, 1086–1099. doi: 10.1016/j.ijmm.2014.08.002

Foucault, M. L., Courvalin, P., and Grillot-Courvalin, C. (2009). Fitness cost of VanA-type vancomycin resistance in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 53, 2354–2359. doi: 10.1128/AAC.01702-08

Frana, T. S., Beahm, A. R., Hanson, B. M., Kinyon, J. M., Layman, L. L., Karriker, L. A., et al. (2013). Isolation and characterization of methicillin-resistant *Staphylococcus aureus* from pork farms and visiting veterinary students. *PLoS One* 8:e53738. doi: 10.1371/journal.pone.0053738

Fuzi, M., Szabo, D., and Csercsik, R. (2017). Double-serine fluoroquinolone resistance mutations advance major international clones and lineages of various multi-drug resistant bacteria. *Front. Microbiol.* 8:2261. doi: 10.3389/fmicb.2017.02261

Geoghegan, J. A., Corrigan, R. M., Gruszka, D. T., Speziale, P., O'Gara, J. P., Potts, J. R., et al. (2010). Role of surface protein SasG in biofilm formation by *Staphylococcus aureus*. *J. Bacteriol.* 192, 5663–5673. doi: 10.1128/JB.00628-10

Ghebremedhin, B., Konig, W., and Konig, B. (2005). Heterogeneity of methicillin-resistant *Staphylococcus aureus* strains at a German university hospital during a 1-year period. *Eur. J. Clin. Microbiol. Infect. Dis.* 24, 388–398. doi: 10.1007/s10096-005-1339-1

Glaser, P., Martins-Simões, P., Villain, A., Barbier, M., Tristan, A., Bouchier, C., et al. (2016). Demography and intercontinental spread of the USA300 community-acquired methicillin-resistant *Staphylococcus aureus* lineage. *mBio* 7:e02183-15. doi: 10.1128/mbio.02183-15

Goerke, C., Pantucek, R., Holtfreter, S., Schulte, B., Zink, M., Grumann, D., et al. (2009). Diversity of prophages in dominant *Staphylococcus aureus* clonal lineages. *J. Bacteriol.* 191, 3462–3468. doi: 10.1128/JB.01804-08

Grundmann, H., Aanensen, D. M., van den Wijngaard, C. C., Spratt, B. G., Harmsen, D., Friedrich, A. W., et al. (2010). Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: a molecular-epidemiological analysis. *PLoS Med.* 7:e1000215. doi: 10.1371/journal.pmed.1000215

Guindon, S., Delsuc, F., Dufayard, J.-F., and Gascuel, O. (2009). "Estimating maximum likelihood phylogenies with PhyML," in *Bioinformatics for DNA*

Sequence Analysis Methods in Molecular Biology, ed. D. Posada (Berlin: Springer Nature).

Gustave, C. A., Tristan, A., Martins-Simões, P., Stegger, M., Benito, Y., Andersen, P. S., et al. (2018). Demographic fluctuation of community-acquired antibiotic-resistant *Staphylococcus aureus* lineages: potential role of flimsy antibiotic exposure. *ISME J.* doi: 10.1038/s41396-018-0110-4 [Epub ahead of print].

Hau, S. J., Bayles, D. O., Alt, D. P., Frana, T. S., and Nicholson, T. L. (2017). Draft genome sequences of 63 swine-associated methicillin-resistant *Staphylococcus aureus* sequence type 5 isolates from the United States. *Genome Announc.* 5:e01081-17. doi: 10.1128/genomeA.01081-17

Hau, S. J., Sun, J., Davies, P. R., Frana, T. S., and Nicholson, T. L. (2015). Comparative prevalence of immune evasion complex genes associated with beta-hemolysin converting bacteriophages in MRSA ST5 isolates from swine, swine facilities, humans with swine contact, and humans with no swine contact. *PLoS One* 10:e0142832. doi: 10.1371/journal.pone.0142832

Holden, M. T., Lindsay, J. A., Corton, C., Quail, M. A., Cockfield, J. D., Pathak, S., et al. (2010). Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). *J. Bacteriol.* 192, 888–892. doi: 10.1128/JB.01255-09

Horváth, A., Dobay, O., Kardos, S., Ghidán, Á., Tóth, Á., Pászti, J., et al. (2012). Varying fitness cost associated with resistance to fluoroquinolones governs clonal dynamic of methicillin-resistant *Staphylococcus aureus*. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 2029–2036. doi: 10.1007/s10096-011-1536-z

Inouye, M., Dashnow, H., Raven, L. A., Schultz, M. B., Pope, B. J., Tomita, T., et al. (2014). SRST2: rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 6:90. doi: 10.1186/s13073-014-0090-6

Kaya, H., Hasman, H., Larsen, J., Stegger, M., Johannesen, T. B., Allesoe, R. L., et al. (2018). SCCmecFinder, a web-based tool for typing of staphylococcal cassette chromosome mec in *Staphylococcus aureus* using whole-genome sequence data. *mSphere* 3, e00612-17. doi: 10.1128/mSphere.00612-17

Knight, G. M., Budd, E. L., Whitney, L., Thornley, A., Al-Ghusein, H., Planche, T., et al. (2012). Shift in dominant hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) clones over time. *J. Antimicrob. Chemother.* 67, 2514–2522. doi: 10.1093/jac/dks245

Kos, V. N., Desjardins, C. A., Griggs, A., Cerqueira, G., Van Tonder, A., Holden, M. T., et al. (2012). Comparative genomics of vancomycin-resistant *Staphylococcus aureus* strains and their positions within the clade most commonly associated with Methicillin-resistant *S. aureus* hospital-acquired infection in the United States. *mBio* 3:e00112-12. doi: 10.1128/mBio.00112-12

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., et al. (2001). Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* 357, 1225–1240. doi: 10.1016/S0140-6736(00)04403-2

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lowder, B. V., Guinane, C. M., Ben Zakour, N. L., Weinert, L. A., Conway-Morris, A., Cartwright, R. A., et al. (2009). Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19545–19550. doi: 10.1073/pnas.0909285106

Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E., and Schneewind, O. (2000). *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5510–5515. doi: 10.1073/pnas.080520697

McDougal, L. K., Steward, C. D., Killgore, G. E., Chaitram, J. M., McAllister, S. K., and Tenover, F. C. (2003). Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. *J. Clin. Microbiol.* 41, 5113–5120. doi: 10.1128/jcm.41.11.5113-5120.2003

Monecke, S., Coombs, G., Shore, A. C., Coleman, D. C., Akpaka, P., Borg, M., et al. (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

Murray, G. G., Wang, F., Harrison, E. M., Paterson, G. K., Mather, A. E., Harris, S. R., et al. (2016). The effect of genetic structure on molecular dating and tests for temporal signal. *Methods Ecol. Evol.* 7, 80–89. doi: 10.1111/2041-210X.12466

Murray, S., Pascoe, B., Meric, G., Mageiros, L., Yahara, K., Hitchings, M. D., et al. (2017). Recombination-mediated host adaptation by avian *Staphylococcus aureus*. *Genome Biol. Evol.* 9, 830–842. doi: 10.1093/gbe/evx037

Mwangi, M. M., Wu, S. W., Zhou, Y., Sieradzki, K., de Lencastre, H., Richardson, P., et al. (2007). Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9451–9456. doi: 10.1073/pnas.0609839104

Nubel, U., Dordel, J., Kurt, K., Strommenger, B., Westh, H., Shukla, S. K., et al. (2010). A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. *PLoS Pathog.* 6:e1000855. doi: 10.1371/journal.ppat.1000855

Omoe, K., Hu, D. L., Ono, H. K., Shimizu, S., Takahashi-Omoe, H., Nakane, A., et al. (2013). Emetic potentials of newly identified staphylococcal enterotoxin-like toxins. *Infect. Immun.* 81, 3627–3631. doi: 10.1128/IAI.00550-13

Omoe, K., Imanishi, K., Hu, D. L., Kato, H., Fugane, Y., Abe, Y., et al. (2005). Characterization of novel staphylococcal enterotoxin-like toxin type P. *Infect. Immun.* 73, 5540–5546. doi: 10.1128/IAI.00550-13

Panesso, D., Planet, P. J., Diaz, L., Hugonnet, J. E., Tran, T. T., Narechania, A., et al. (2015). Methicillin-susceptible, vancomycin-resistant *Staphylococcus aureus*, Brazil. *Emerg. Infect. Dis.* 21, 1844–1848. doi: 10.3201/eid2110.141914

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26, 419–420. doi: 10.1093/bioinformatics/btp696

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412

Planet, P. J., Diaz, L., Kolokotronis, S.-O., Narechania, A., Reyes, J., Xing, G., et al. (2015). Parallel epidemics of community-associated methicillin-resistant *Staphylococcus aureus* USA300 infection in North and South America. *J. Infect. Dis.* 212, 1874–1882. doi: 10.1093/infdis/jiv320

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Roberts, R. B., de Lencastre, A., Eisner, W., Severina, E. P., Shopsin, B., Kreiswirth, B. N., et al. (1998). Molecular epidemiology of methicillin-resistant Staphylococcus aureus in 12 New York hospitals. MRSA Collaborative Study Group. *J. Infect. Dis.* 178, 164–171. doi: 10.1086/515610

Robinson, D. A., and Enright, M. C. (2004). Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 10, 92–97. doi: 10.1111/j.1469-0691.2004.00768.x

Rodriguez-Noriega, E., Seas, C., Guzman-Blanco, M., Mejia, C., Alvarez, C., Bavestrello, L., et al. (2010). Evolution of methicillin-resistant *Staphylococcus aureus* clones in Latin America. *Int. J. Infect. Dis.* 14, e560–e566. doi: 10.1016/j.ijid.2009.08.018

Rolain, J. M., François, P., Hernandez, D., Bittar, F., Richet, H., Fournous, G., et al. (2009). Genomic analysis of an emerging multiresistant *Staphylococcus aureus* strain rapidly spreading in cystic fibrosis patients revealed the presence of an antibiotic inducible bacteriophage. *Biol. Direct* 4:1. doi: 10.1186/1745-6150-4-1

Sansevere, E. A., Luo, X., Park, J. Y., Yoon, S., Seo, K. S., and Robinson, D. A. (2017). Transposase-mediated excision, conjugative transfer, and diversity of ICE*6013* elements in *Staphylococcus aureus*. *J. Bacteriol.* 199:e00629-16. doi: 10.1128/JB.00629-16

Sansevere, E. A., and Robinson, D. A. (2017). Staphylococci on ICE: overlooked agents of horizontal gene transfer. *Mob. Genet. Elements* 7, 1–10. doi: 10.1080/2159256X.2017.1368433

Schulte, B., Bierbaum, G., Pohl, K., Goerke, C., and Wolz, C. (2013). Diversification of clonal complex 5 methicillin-resistant *Staphylococcus aureus* strains (Rhine-Hesse clone) within Germany. *J. Clin. Microbiol.* 51, 212–216. doi: 10.1128/JCM.01967-12

Smyth, D. S., McDougal, L. K., Gran, F. W., Manoharan, A., Enright, M. C., Song, J. H., et al. (2010). Population structure of a hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. *PLoS One* 5:e8582. doi: 10.1371/journal.pone.0008582

Sola, C., Cortes, P., Saka, H. A., Vindel, A., and Bocco, J. L. (2006). Evolution and molecular characterization of methicillin-resistant *Staphylococcus aureus* epidemic and sporadic clones in Cordoba, Argentina. *J. Clin. Microbiol.* 44, 192–200. doi: 10.1128/JCM.44.1.192-200.2006

Sola, C., Gribaudo, G., Vindel, A., Patrito, L., Bocco, J. L., and Cordoba MRSA Collaborative Study Group (2002). Identification of a novel methicillin-resistant *Staphylococcus aureus* epidemic clone in Cordoba, Argentina, involved in nosocomial infections. *J. Clin. Microbiol.* 40, 1427–1435. doi: 10.1128/JCM.40.4.1427-1435.2002

Sola, C., Saka, H. A., Cordoba MRSA Collaborative Study Group Vindel, A., and Bocco, J. L. (2008). Emergence and dissemination of a community-associated methicillin-resistant Panton-Valentine leucocidin-positive *Staphylococcus aureus* clone sharing the sequence type 5 lineage with the most prevalent nosocomial clone in the same region of Argentina. *J. Clin. Microbiol.* 46, 1826–1831. doi: 10.1128/JCM.01949-07

Strommenger, B., Bartels, M. D., Kurt, K., Layer, F., Rohde, S. M., Boye, K., et al. (2014). Evolution of methicillin-resistant *Staphylococcus aureus* towards increasing resistance. *J. Antimicrob. Chemother.* 69, 616–622. doi: 10.1093/jac/dkt413

Tickler, I. A., Goering, R. V., Mediavilla, J. R., Kreiswirth, B. N., Tenover, F. C., and Consortium, H. A. I. (2017). Continued expansion of USA300-like methicillin-resistant *Staphylococcus aureus* (MRSA) among hospitalized patients in the United States. *Diagn. Microbiol. Infect. Dis.* 88, 342–347. doi: 10.1016/j.diagmicrobio.2017.04.016

Uhlemann, A. C., Dordel, J., Knox, J. R., Raven, K. E., Parkhill, J., Holden, M. T. G., et al. (2014). Molecular tracing of the emergence, diversification, and transmission of *S. aureus* sequence type 8 in a New York community. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6738–6743. doi: 10.1073/pnas.1401006111

van Wamel, W. J., Rooijakkers, S. H., Ruyken, M., van Kessel, K. P., and van Strijp, J. A. (2006). The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *Staphylococcus aureus* are located on beta-hemolysin-converting bacteriophages. *J. Bacteriol.* 188, 1310–1315. doi: 10.1128/JB.188.4.1310-1315.2006

Velazquez-Meza, M. E., Aires de Sousa, M., Echaniz-Aviles, G., Solorzano-Santos, F., Miranda-Novales, G., Silva-Sanchez, J., et al. (2004). Surveillance of methicillin-resistant *Staphylococcus aureus* in a pediatric hospital in Mexico City during a 7-year period (1997 to 2003): clonal evolution and impact of infection control. *J. Clin. Microbiol.* 42, 3877–3880. doi: 10.1128/JCM.42.8.3877-3880.2004

Vergara-Irigaray, M., Valle, J., Merino, N., Latasa, C., García, B., Ruiz de Los Mozos, I., et al. (2009). Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. *Infect. Immun.* 77, 3978–3991. doi: 10.1128/IAI.00616-09

World Health Organization (2014). *Antimicrobial Resistance: Global Report on Surveillance.* Available at: http://www.who.int/drugresistance/documents/surveillancereport/en/ [accessed May 1, 2018].

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T. Y. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210X.12628

Molecular Typing of ST239-MRSA-III From Diverse Geographic Locations and the Evolution of the *SCCmec* III Element During Its Intercontinental Spread

Stefan Monecke [1,2,3*], Peter Slickers [1,2], Darius Gawlik [1,2], Elke Müller [1,2], Annett Reissig [1,2], Antje Ruppelt-Lorz [3], Patrick E. Akpaka [4], Dirk Bandt [5], Michele Bes [6], Samar S. Boswihi [7], David C. Coleman [8], Geoffrey W. Coombs [9], Olivia S. Dorneanu [10], Vladimir V. Gostev [11], Margaret Ip [12], Bushra Jamil [13,14], Lutz Jatzwauk [15], Marco Narvaez [15], Rashida Roberts [4], Abiola Senok [16], Anna C. Shore [8], Sergey V. Sidorenko [12], Leila Skakni [17], Ali M. Somily [18], Muhammad Ali Syed [19], Alexander Thürmer [3], Edet E. Udo [8], Teodora Vremerǎ [10], Jeannete Zurita [20,21] and Ralf Ehricht [1,2]

[1] Abbott (Alere Technologies GmbH), Jena, Germany, [2] InfectoGnostics Research Campus Jena, Jena, Germany, [3] Medical Faculty "Carl Gustav Carus", Institute for Medical Microbiology and Hygiene, Technische Universität Dresden, Dresden, Germany, [4] Department of Paraclinical Sciences, The University of the West Indies, St. Augustine, Trinidad and Tobago, [5] Instituts für Labordiagnostik, Mikrobiologie und Krankenhaushygiene, Oberlausitz-Kliniken, Bautzen, Germany, [6] Centre National de Référence des Staphylocoques, Institut des Agents Infectieux, Hospices Civils de Lyon, Lyon, France, [7] Microbiology Department, Faculty of Medicine, Kuwait University, Kuwait City, Kuwait, [8] Microbiology Research Unit, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin, Ireland, [9] School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, Australia, [10] Microbiology Unit, Department of Preventive and Interdisciplinary Medicine, University of Medicine & Pharmacy "Grigore T Popa", Iaşi, Romania, [11] Pediatric Research and Clinical Center for Infectious Diseases, Saint Petersburg, Russia, [12] Department of Microbiology, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, [13] Department of Biosciences, COMSATS Institute of Information Technology, Islamabad, Pakistan, [14] Department of Biogenetics, National University of Medical Sciences, Rawalpindi, Pakistan, [15] Department of Hospital Infection Control, Dresden University Hospital, Dresden, Germany, [16] Department of Basic Medical Sciences, College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates, [17] Molecular Pathology Laboratory, King Fahad Medical City, Riyadh, Saudi Arabia, [18] Department of Pathology and Laboratory Medicine, College of Medicine, King Saud University and King Saud University Medical City, Riyadh, Saudi Arabia, [19] Department of Microbiology, University of Haripur, Haripur, Pakistan, [20] Facultad de Medicina, Pontificia Universidad Catolica del Ecuador, Quito, Ecuador, [21] Zurita & Zurita Laboratorios, Unidad de Investigaciones en Biomedicina, Quito, Ecuador

ST239-MRSA-III is probably the oldest truly pandemic MRSA strain, circulating in many countries since the 1970s. It is still frequently isolated in some parts of the world although it has been replaced by other MRSA strains in, e.g., most of Europe. Previous genotyping work (Harris et al., 2010; Castillo-Ramírez et al., 2012) suggested a split in geographically defined clades. In the present study, a collection of 184 ST239-MRSA-III isolates, mainly from countries not covered by the previous studies were characterized using two DNA microarrays (i) targeting an extensive range of typing markers, virulence and resistance genes and (ii) a SCCmec subtyping array. Thirty additional isolates underwent whole-genome sequencing (WGS) and, together with published WGS data for 215 ST239-MRSA-III isolates, were analyzed using in-silico analysis for comparison with the microarray data and with special regard to variation within SCCmec elements. This permitted the assignment of isolates and sequences to 39 different SCCmec III subtypes, and to three major and several minor clades. One clade, characterized by the integration

of a transposon into *nsaB* and by the loss of *fnbB* and *splE* was detected among isolates from Turkey, Romania and other Eastern European countries, Russia, Pakistan, and (mainly Northern) China. Another clade, harboring *sasX/sesI* is widespread in South-East Asia including China/Hong Kong, and surprisingly also in Trinidad & Tobago. A third, related, but *sasX/sesI*-negative clade occurs not only in Latin America but also in Russia and in the Middle East from where it apparently originated and from where it also was transferred to Ireland. Minor clades exist or existed in Western Europe and Greece, in Portugal, in Australia and New Zealand as well as in the Middle East. Isolates from countries where this strain is not epidemic (such as Germany) frequently are associated with foreign travel and/or hospitalization abroad. The wide dissemination of this strain and the fact that it was able to cause a hospital-borne pandemic that lasted nearly 50 years emphasizes the need for stringent infection prevention and control and admission screening.

INTRODUCTION

Staphylococcus aureus is a bacterial species that colonizes the skin and mucous membranes of a high percentage of the human population (van Belkum et al., 2009) and several animal species. It can cause localized infections, such as skin and soft tissue infections (SSTIs), bone, joint and implant infections or pneumonia as well as sepsis and toxicoses including toxic shock syndrome. Resistance toward antibiotics in *S. aureus* is a highly relevant issue. Methicillin resistance and resistance to most beta-lactams is due to the production of modified penicillin-binding proteins encoded by *mec* genes. The *mecA/mecC* genes are located on large, complex and potentially mobile staphylococcal cassette chromosome (SCC*mec*) elements (while *mecB* was observed on a plasmid; Becker et al., 2018). SCC*mec* elements additionally encode regulatory elements and, variably, genes encoding resistance to other antimicrobials, such as aminoglycosides, macrolides, tetracyclines, fusidic acid and to heavy metals (Oliveira et al., 2000; Ito et al., 2001). The comparatively older SCC*mec* types I, II, and III are typically restricted to MRSA strains involved in healthcare-associated infections (HCA-MRSA).

The HCA-MRSA, sequence type (ST) 239 MRSA, as designated by multilocus sequence typing (MLST), is of special interest. As ST240 and ST241 are single locus variants of ST239 (which differ only by mutations in MLST marker genes *pta* or *yqiI*), these STs are here discussed together as clonal complex (CC) 239. CC239 harboring SCC*mec* type III have been designated various names in different geographic regions including "Wiener Epidemiestamm" (Vienna Epidemic Strain), the Hungarian Clone, UK-EMRSA−1,−4,−7,−9, or−11, Irish Phenotype III, Irish AR01,−09,−44, and−23, the Brazilian Clone, Australian Epidemic MRSA−2 and−3 as well as Canadian MRSA−3 or−6. CC239-MRSA-III is probably the oldest truly pandemic MRSA strain. In contrast to other early MRSA strains, it is still common and widespread, at least in some parts of the world.

CC239-MRSA-III has been reported in many European countries including the UK (Edgeworth et al., 2007), Ireland (Humphreys et al., 1990; Shore et al., 2005), Spain (Cuevas et al., 2007), Portugal (Smyth et al., 2010), Italy (Campanile et al., 2009), Malta (Scicluna et al., 2010), Croatia (Budimir et al., 2009), Germany (Witte et al., 1997; Albrecht et al., 2011), Austria (Krziwanek et al., 2008), the Czech Republic (Melter et al., 2003), Hungary (Conceicao et al., 2007), and Greece (Aires de Sousa et al., 2003a). It is also frequently observed in Romania (Cirlan et al., 2005; Chen et al., 2010; Monecke et al., 2014a) and Russia (Afanas'ev et al., 2010; Baranovich et al., 2010; Yamamoto et al., 2012; Gostev et al., 2017). CC239-MRSA-III is common to abundant in Mediterranean and Middle Eastern countries such as Greece (Aires de Sousa et al., 2003a), Turkey (Alp et al., 2009; Tekeli et al., 2016), Egypt (El-baz et al., 2017), Morocco, Tunisia, Algeria (Abdulgader et al., 2015), Iran (Fatholahzadeh et al., 2009), Saudi Arabia (Cirlan et al., 2005; Monecke et al., 2012c; Senok et al., 2016), Abu Dhabi (Weber et al., 2010), and Kuwait (Boswihi et al., 2016). The strain's presence has been reported in China (Chen et al., 2014a,b) including Hong Kong (Ip et al., 2005), Taiwan (Aires de Sousa et al., 2003b; Takano et al., 2007), Singapore (Hsu et al., 2007), Malaysia (Ghaznavi-Rad et al., 2010), Mongolia (Orth et al., 2006), Pakistan (Shabir et al., 2010; Zafar et al., 2011; Arfat, 2013; Jamil et al., 2017), India (D'Souza et al., 2010; Neetu and Murugan, 2016), South Korea (Cha et al., 2005; Peck et al., 2009), Laos (Yeap et al., 2017), and Thailand (Smyth et al., 2010). There are reports from several African countries including Ghana, Kenya, Niger, Nigeria, Senegal, and South Africa (Jansen van Rensburg et al., 2011; Abdulgader et al., 2015). In Eastern Australia and New Zealand it was a common cause of HCA infection, and it has been reported in association with large MRSA outbreaks (Coombs et al., 2004; Howden et al., 2010). In the western hemisphere, CC239-MRSA-III has been reported from the United States (Schaefler et al., 1981), Brazil (Vivoni et al., 2006), Paraguay (Mayor et al., 2007), Ecuador (Zurita et al., 2016), and Trinidad & Tobago where it recently

was still the predominant MRSA strain (Akpaka et al., 2007; Monecke et al., 2012b, 2014b).

In terms of genetic structure, ST239 is an interesting clone, with six of the seven MLST housekeeping genes being identical to ST8 (Robinson and Enright, 2004a,b). However ST239 as well as ST240 and ST241 harbor a very different *arcC*-allele (*arcC*-2, rather than *arcC*-3). The difference in the *arcC*-allele and the presence of some other features that differentiate ST239 from canonical ST8 (including the presence of the collagen adhesion gene *cna*, capsule type 5 and RIDOM *spa* types t030 or t037), indicate the integration of a large fragment of genomic DNA of clonal complex (CC) 30 origin into a CC8 chromosome (Robinson and Enright, 2004b; Holden et al., 2009). The fragment consists of ~635,000 base pairs, which is ~20% of a *S. aureus* genome. The mechanism of integration is yet not known.

Another interesting recent observation has been the discovery of *sasX/sesI* in CC239-MRSA, a virulence factor thought to have a key role in nasal colonization, pathogenesis of lung disease, and abscess formation (Li et al., 2012). The *sasX* gene is located on a 127 kb lysogenic prophage phiSPbeta (Li et al., 2012) and it encodes the surface-anchored protein X, an LPxTG motif surface-anchored protein, and does not have orthologues in any of the other sequenced *S. aureus* genomes. A highly similar gene, *sesI*, is present in the *S. epidermidis* phiSPbeta region and has also been identified in other coagulase-negative staphylococci such as *S. capitis* (GenBank JGYJ) and *S. cohnii* (GenBank LATU and LATV).

With the rise of Next Generation Sequencing (NGS) technologies, Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012) sequenced a large collection of CC239-MRSA-III from very diverse geographic origins. They suggested a phylogenetic framework in which isolates of CC239-MRSA-III clustered in several major "clades" and a couple of isolated branches. The clades are largely associated with geographic background and thus were referred to as "European", "Latin American", "Turkish", and "Asian" clades.

In the present study, a collection of CC239-MRSA-III isolates was characterized, primarily using previously published DNA microarray technology targeting typing markers, virulence and resistance genes and SCC*mec* subtypes (Monecke et al., 2008b, 2011, 2016). Published whole-genome sequence data, in particular those by Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012), were re-analyzed with regard to the presence or absence of the marker genes as used experimentally for array hybridization and with regard to variation within SCC*mec* elements. A comparison between our strain collection, mainly from countries that were not covered by the Harris and Castillo-Ramírez' work, to published genome sequences was then performed in order to see if and how they fit into their proposed phylogenetic framework and to look for epidemiological connections and suitable marker genes.

MATERIALS AND METHODS

Isolates

In total, 214 clinical or screening isolates were included in the present study. These isolates originated from hospitals in Ireland,

Germany, Romania, Kuwait, Saudi Arabia, Russia, Pakistan, China/Hong Kong, Australia, Trinidad&Tobago, and Ecuador as well as some reference strains (see **Supplemental Table 1** and below). Some of the isolates were a convenience sample from previous studies (Monecke et al., 2008a, 2011, 2012b,c, 2014a,b, 2016; Albrecht et al., 2011; Boswihi et al., 2016; Senok et al., 2016; Zurita et al., 2016; Gostev et al., 2017; Jamil et al., 2017). Others came from ongoing routine diagnostics, outbreak investigations or typing tasks performed by the authors and have not been published previously. Only one isolate per patient was included. Isolates were stored frozen using cryobank tubes (Microbank, Pro-Lab Diagnostics, Richmond Hill, Canada) at −80°C. Isolates were routinely cultured on Columbia blood agar plates, and DNA preparation was performed as previously described (Monecke et al., 2008b, 2011).

DNA Microarrays for SCC*mec* Typing and Subtyping

Two microarrays were used in the study, and were applied to 184 of the isolates investigated herein (another 30 were directly subjected to sequencing; see below). Both arrays have previously been described including probe and primer sequences, details of DNA extraction, labeling, amplification, hybridization protocols as well as data analysis and interpretation. The first microarray (Monecke et al., 2008b, 2011) detects genes associated with antibiotic resistance and virulence as well as a multitude of genes that can be used for typing purposes, such as genes related to *agr* or capsule types, *set/ssl* genes and genes coding for adhesion factors. In addition to the detection of individual genes, the array also allows the assignment of isolates to MLST CCs, to known epidemic strains and to SCC*mec* types. The second microarray (Monecke et al., 2016) was designed to subtype SCC*mec* elements. It also includes probes for *sasX/sesI* and for heavy metal resistance genes. Furthermore, it included a set of probes termed "SCCterm" followed by a number (see **Table 1**) which were designed to recognize intergenic regions alternative to *dcs*, between *orfX* and the first codon on the SCC*mec* element.

All markers of relevance for SCC*mec* subtyping are listed in the **Supplemental Table 1**. Markers that were detected among isolates in the present study are listed, and presented in more detail, in **Table 1**.

DNA Microarrays for *mecA* Subtyping

A third assay was used to identify and categorize alleles and variants of the *mecA/C* gene as previously described (Monecke et al., 2012a). While there are a multitude of *mecA* variants (also named *mecA*1) in staphylococcal species other than *S. aureus*, only four different *mecA/C* alleles are of relevance in *S. aureus*/MRSA that also differ in amino acid sequences encoded, i.e., (named with respect to representative GenBank entries), $mecA_{(CP000046)}$ (as in the CC8-MRSA-I strain COL), $mecA_{(BA000018)}$ (as in the CC5-MRSA-II strain N315), $mecA_{(GQ902038)}$ (as in the CC398-MRSA-VT strain UMCG-M4) and *mecC* (GenBank NG_047955.1). This array was applied to 30 isolates representing at least one isolate per SCC*mec* subtype.

TABLE 1 | SCCmec-associated genes and other relevant markers that were targeted by probes used for this study.

Gene/Marker	Gene product/function	Comments	Reference sequence	Reference for probes and primers
aacA-aphD	Bifunctional enzyme Aac/Aph, gentamicin and tobramycin resistance	–	GU565967.1: 24980...26419	Monecke et al., 2008b, 2011
aadD	Aminoglycoside adenyltransferase, tobramycin resistance	–	BA000017.4: 40819...41580	Monecke et al., 2008b, 2011
aadE	Streptomycin aminoglycoside 6-adenyltransferase	Part of a gene cluster comprising of aadE, sat, aphA3	AB300568.1:1...909	Probe: AB300568.1 [643:669] Primer: AB300568.1 [714:733:r]
ant9	Streptomycin 3'-adenylyltransferase	–	AJ810120.1: 14339...15121	Probe: AJ810120.1 [14651:14678] Primer: AJ810120.1 [14681:14700:r]
aphA3	3'/5'-aminoglycoside phosphotransferase, neo- and kanamycin resistance	Part of a gene cluster comprising of sat, aphA3 and variably aadE	U51474.1: 957...1751	Monecke et al., 2008b, 2011
arcA-SCC	Arginine deiminase	Part of ACME I and ACME II clusters, that occurs alone or in combination with SCCmec elements	CP000255.1: 73113...74348	Monecke et al., 2008b, 2011
arcB-SCC	Ornithine carbamoyltransferase	Part of ACME I and ACME II clusters, that occurs alone or in combination with SCCmec elements	CP000255.1: 69839...70837	Monecke et al., 2008b, 2011
arcC-SCC	Carbamate kinase	Part of ACME I and ACME II clusters, that occurs alone or in combination with SCCmec elements	CP000255.1: 68890...69819	Monecke et al., 2008b, 2011
arcD-SCC	Arginine/ornithine antiporter	Part of ACME I and ACME II clusters, that occurs alone or in combination with SCCmec elements	CP000255.1: 71606...73027	Monecke et al., 2008b, 2011
arsC-SCC	Arsenate reductase	Probe was designed for SCC-borne allele(s) of that gene, as for instance in JCSC6943, JCSC6945	AB505628.1: 31467...31868	Monecke et al., 2016
blaZ	Penicillinase	Plasmid-borne penicillinase. This does NOT include the allele from SCCmec XI which is of no relevance for the present study.	M15526.1: 140...985	Monecke et al., 2008b, 2011
ble	Bleomycin resistance protein	–	AF181950.1: 4163...4567	Probe: AF181950.1 [4246:4270] Primer: AF181950.1 [4276:4296:r]
cadD (R35)	Cadmium transport protein D	Probe was designed for SCC-borne allele(s) of that gene, as for instance in strain R35, GenBank L10909.1 or strain 85/2082, GenBank AB037671.1.	L10909.1: 5577...6194	Monecke et al., 2016
cadX	Putative regulator of cadmium efflux	Can be part of SCCmec elements (as in JCSC6943; GenBank AB505628.1) but **Table 3** refers to the plasmid-borne variant	BX571858.1: 8646...8993	Probes: BX571858.1 [8797:8827] and [8676:8705] Primers: BX571858.1 [8855:8874:r] and [8762:8783:r]
cat	Chloramphenicol acetyltransferase	Located on different plasmids	X02529.1: 2267...2914 AY355285.1: 1000...1647 U35036.1: 1167...1814	Monecke et al., 2008b, 2011

(Continued)

TABLE 1 | Continued

Gene/Marker	Gene product/function	Comments	Reference sequence	Reference for probes and primers
ccrA-3	Cassette chromosome recombinase A, type 3	Cassette chromosome recombinase A allele found in SCCmec III elements	AB037671.1: 5430...6776	Monecke et al., 2008b, 2011
ccrA-4	Cassette chromosome recombinase A, type 4	Cassette chromosome recombinase A allele found in SCCmec VI and SCCmec VIII elements	AF411935.3: 7849...9210	Monecke et al., 2008b, 2011
ccrAA	Cassette chromosome recombinase AA	Gene for a hypothetical protein accompanying ccrC. The allele in CC239 is poorly (ambiguously) recognized by the probe herein, but sequence analysis shows it presence. Hence, it is indicated in **Tables 2a,b** in brackets.	AB121219.1: 14264...15907 AM292304.1: 5654...7273	Monecke et al., 2008b, 2011
ccrB-3	Cassette chromosome recombinase B, type 3	Cassette chromosome recombinase B allele found in SCCmec III elements	AB037671.1: 6797...8425	Monecke et al., 2008b, 2011
ccrB-4	Cassette chromosome recombinase B, type 4	Cassette chromosome recombinase B allele found in SCCmec VI and SCCmec VIII elements	AE015929.1: 58592...60220	Monecke et al., 2008b, 2011
ccrC	Cassette chromosome recombinase C	Cassette chromosome recombinase C allele found in in SCCmec V, SCCmec VT and SCCmec VII elements. Sequence analysis (but not array hybridization) allows split into two alleles.	ccrC(TSGH17): AB512767.1: 25157...26836 ccrC(PM1): AB462393.1: 7514...9190	Monecke et al., 2008b, 2011
chp	Chemotaxis-inhibiting protein (CHIPS)	Situated on a hlb-integrating phage	BX571856.1: 2126786...2127235	Monecke et al., 2008b, 2011
cstB-SCC1 (Q2G1R6)	CsoR-like sulfur transferase-regulated genes B/metallo-beta-lactamase superfamily protein. Pseudogene containing two stop codons	Subtyping SCCmec II (usually present, but absent in Irish SCCmec II variants A to E) and III (usually present, but absent in CMFT492 GenBank HF569112.1). Also present in SCCmec VIII and irregular elements such as Staphylococcus fleurettii GenBank AB546266	BA000017.4: 50957...52291	Monecke et al., 2016
cstB-SCC2 (Q2G1R6)	CsoR-like sulfur transferase-regulated genes B/metallo-beta-lactamase superfamily protein.	Present in SCCmec IVa (truncated) and SCCmec X, variably present in SCCmec I and VT	CP000046.1: 53428...54756	Monecke et al., 2016
czrC	Cadmium and zinc resistance gene C, heavy metal translocating P-type ATPase	Frequently associated with SCCmec elements from livestock MRSA	AE015929.1: 64066...66000	Monecke et al., 2016
dfrA	Dihydrofolate reductase mediating trimethoprim resistance	–	AF051916.1: 2823...3308	Monecke et al., 2008b, 2011
dfrG	Dihydrofolate reductase mediating trimethoprim resistance	–	AB205645.1: 1013...1510	Probes: AB205645.1 [1039:1068] and [1142:1167] Primers: AB205645.1[1082:1098:r] and [1172:1189:r]
D1GU38	Putative protein	Subtyping SCCmec III, identification of SCCmec VT, SCCmec ZH47, SCCmec VII because of an association with (additional/second) ccrC copies	FN433596.1: 34888...35751	Monecke et al., 2016
D1GU55	Putative membrane protein	Subtyping SCCmec III, additional marker for SCCmec VII	FN433596.1: 52909...53208	Monecke et al., 2016
erm(A)	rRNA adenine N-6-methyltransferase, erythromycin/clindamycin resistance	–	D86934.2: 35074...35805	Monecke et al., 2008b, 2011
erm(C)	rRNA adenine N-6-methyltransferase, erythromycin/clindamycin resistance	–	V01278.1: 2004...2738	Monecke et al., 2008b, 2011

(Continued)

TABLE 1 | Continued

Gene/Marker	Gene product/function	Comments	Reference sequence	Reference for probes and primers
fnbA	Fibronectin-binding protein A	–	CP000046.1: 2569440...2572496	Monecke et al., 2008b, 2011
fnbB	Fibronectin-binding protein B	–	CP000046.1: 2565936...2568758	Monecke et al., 2008b, 2011
hla	Haemolysin alpha, sphingomyelin phosphodiesterase	–	X13404.1: 752...1744	Monecke et al., 2008b, 2011
mecA	Modified penicillin-binding protein (PBP2a)	Modified penicillin-binding protein (PBP2a) causing oxacillin/methicillin resistance and thus defining MRSA	BA000017.4: 44992...46998	Monecke et al., 2008b, 2011
mecI	Methicillin-resistance regulatory protein	Present in SCCmec II (although absent from Irish SCCmec II variants C and E), SCCmec III, SCCmec VIII	BA000017.4: 48855...49226	Monecke et al., 2008b, 2011
mecR1	Methicillin resistance operon repressor 1; Un-truncated sequence in SCCmec II, III, VIII. Truncated mecR1 is present in SCCmec I, IV, V, VI, VII. Absent from SCCmec V, IX, X.	Two probes are utilized. Both yield signals in SCCmec II, SCCmec III, SCCmec VIII while only the Δ mecR1 probe gives a signal for the truncated mecR1 of SCCmec I, IV, VI, VII. Some "European Clade" strains related to 85/2082, GenBank AB037671.1 have another truncation resulting in a signal for the probe associated with untruncated mecR1 but an absence of a signal for the Δ mecR1 probe.	BA000017.4: 47098...48855	Monecke et al., 2008b, 2011
merA	Mercury reductase	Part of a mercury resistance operon that is plasmid-borne, although the plasmid can be integrated into SCCmec elements (for instance, in strains 85/2082, GenBank AB037671.1 or TW20, GenBank FN433596.1)	AB037671.1: 38289...39932	Monecke et al., 2008b, 2011
merB	Alkylmercury lyase	Part of a mercury resistance operon that is plasmid-borne, although the plasmid can be integrated into SCCmec elements (for instance, in strain TW20, GenBank FN433596.1)	AB037671.1: 37557...38207	Monecke et al., 2008b, 2011
mupA	Mupirocin resistance protein, alternative isoleucyl-tRNA synthase	-	X75439.1:..477...3551	Monecke et al., 2008b, 2011
mvaS-SCC	Truncated 3-hydroxy-3-methylglutaryl CoA synthase	Subtyping SCCmec I, II, IV, V	BA000033.2: 37179...37531	Monecke et al., 2016
PSM-mec	Phenol soluble modulin from SCCmec	Present in SCCmec II (although absent from Irish SCCmec II variants C and E), SCCmec III, SCCmec VIII	BA000017.4: 49311...49379	(Monecke et al., 2016)
Q4LAG7	Putative protein located within SCCmec type V/SCCfus elements	Identification of SCCmec V/VT elements and of SCCfus elements	AM990992.1: 50512...50940 (V/VT) BX571857.1: 55452...55880 (fus)	Monecke et al., 2016
Q8CU82	Putative protein	Present in some SCCmec/fus composite elements such as e.g., CMFT120, GenBank HF569094.1 and CMFT2, GenBank HF569101.1	AE015929.1: 32604...32786	Monecke et al., 2016
Q933A2	Putative ADP-ribosyltransferase	Subtyping SCCmec III and SCCmec IX	FN433596.1: 101805...102377	Monecke et al., 2016
Q93IB7	LytTR domain DNA-binding regulator	Subtyping SCCmec III (present in, e.g., TW20 GenBank FN433596.1, but absent in, e.g., Bmb9393 GenBank CP005288.1) and IV (usually absent, but present in, e.g., CMFT503 GenBank HF569113.1)	FN433596.1: 67873...68115	Monecke et al., 2016

(Continued)

TABLE 1 | Continued

Gene/Marker	Gene product/function	Comments	Reference sequence	Reference for probes and primers
Q9S0M4	Putative protein	Subtyping SCCmec I, SCCmec/ACME composites and SCCmec from WA40	JQ746621.1: 10406...11456	Monecke et al., 2016
Q9XB68-dcs	Located at the terminus of SCCmec directly next to orfX	This locus comprises the downstream constant segment (dcs) that in turn comprises a copy of the SCC direct repeat DR_SCC (AGAAGCTTATCATAAGTAA).	dcs: CP000046.1: 34192...34371 Q9XB68: CP000046.1: 34372...35667	Monecke et al., 2008b, 2011
qacA	Quaternary ammonium compound resistance protein A	–	AF535086.1: 29...1573	Monecke et al., 2008b, 2011
qacC	Quaternary ammonium compound resistance protein C	–	AY121858.1: 261...584	Monecke et al., 2008b, 2011
sak	Staphylokinase	Situated on a hlb-integrating phage	AF424783.1: 38467...38958	Monecke et al., 2008b, 2011
sat	Streptothricin-acetyltransferase	Part of a gene cluster comprising of sat, aphA3 and variably aadE	CP000028.1: 8575...9105	Monecke et al., 2008b, 2011
sasX/sesI	Surface-anchored protein X	Located on 127 kb prophage phiSPbeta	FN433596.1: 2307955...2308569	Probes: FN433596.1 [2308472:2308498:r] and [2308232:2308256:r] Primers: FN433596.1 [2308444:2308466] and [2308203:2308220]
SCC terminus 01	SCC integration site alternate to dcs	–	GU235983.1: 488...808	Monecke et al., 2016
SCC terminus 02	SCC integration site alternate to dcs	–	FN433596.1: 34140...34456	Monecke et al., 2016
SCC terminus 03	SCC integration site alternate to dcs	–	FR753166.1: 481...568	Monecke et al., 2016
SCC terminus 05	SCC integration site alternate to dcs	–	AB425427.1: 606...1027	Monecke et al., 2016
SCC terminus 07	SCC integration site alternate to dcs	–	GU122149.1: 119...222	Monecke et al., 2016
scn	Staphylococcal complement inhibitor (SCIN)	Situated on a hlb-integrating phage	AF424783.1: 41170...41520	Monecke et al., 2008b, 2011
sea	Enterotoxin A	Situated on a hlb-integrating phage	M18970.1: 1...774	Monecke et al., 2008b, 2011
sea(N315)/sep	Enterotoxin A allele from strain N315 (GenBank BA000018.3), also known as Enterotoxin P	Situated on a hlb-integrating phage	BA000018.3: 2011380...2012153	Monecke et al., 2008b, 2011
sek	Enterotoxin K	–	CP000046.1: 904800...905528	Monecke et al., 2008b, 2011
seq	Enterotoxin Q	–	CP000046.1: 905552...906280	Monecke et al., 2008b, 2011

(Continued)

TABLE 1 | Continued

Gene/Marker	Gene product/function	Comments	Reference sequence	Reference for probes and primers
splA	Serine protease A	–	AF271715.1: 1001...1708	Monecke et al., 2008b, 2011
splB	Serine protease B	–	AF271715.1: 1833...2555	Monecke et al., 2008b, 2011
splE	Serine protease E	–	AF271715.1: 4330...5046	Monecke et al., 2008b, 2011
tet(K)	Tetracycline efflux protein variant K	–	U38656.1: 436...1815	Monecke et al., 2008b, 2011
tet(M)	Ribosomal protection tetracycline resistance protein M	–	M21136.1: 458..2377	Monecke et al., 2008b, 2011
tst1	Toxic shock syndrome toxin 1	–	BA000017.4: 2137509...2138213	Monecke et al., 2008b, 2011
ugpQ	Glycerophosphoryl diester phosphodiesterase	Accompanies mecA in all S. aureus SCCmec sequences except SCCmec IV A from CN1 GenBank CP003979.1	BA000018.3: 43717...44460	Monecke et al., 2008b, 2011
xylR/mecR2	Methicillin resistance operon repressor 2, Homolog of xylose repressor	Located next to mec operon downstream of mecI (not present if mecI is truncated). Present in SCCmec II variants C and E), SCCmec III, SCCmec VIII	BA000018.3: 49738...50882	Monecke et al., 2008b, 2011
		(although absent from Irish SCCmec II variants C and E)		

SCCmec Subtypes and Nomenclature

The guidelines of the International Working Group (IWG-SCC, 2009) were used for the assignment of Roman numerals to SCCmec types defined by the class of mec gene complex and type of cassette chromosome recombinase (ccr) genes. As proposed by Shore and Coleman (2013) we named elements lacking ccr recombinase genes "pseudoSCCmec elements." Composite elements are indicated by listing relevant components in square brackets. Heavy metal resistance genotypes were described by adding chemical symbols rather than individual gene designations e.g., SCC [mec III+Cd/Hg+ccrC]).

Genes aadD, aacA-aphD, ant9, ble, and erm(A) as well as tet genes were not included in the analysis of the SCCmec III subtypes. Although these genes may be situated on SCCmec elements, they also may be found on plasmids or other mobile genetic elements at various locations [see below for erm(A) and ant9]. Neither array hybridization nor those NGS technologies that yield a high number of short contigs can provide reliable information on the actual localizations of genes, or whether a plasmid [such as pT181/tet(K)] was free or integrated into the genome. Consequently, we did not differentiate between SCCmec III and IIIA (Vandenesch et al., 2003).

In contrast, the mercury resistance operon was included into the analysis of the SCCmec III subtypes. In CC239 this operon is part of a composite SCC element, although it can indeed be found outside of SCC elements (e.g., GenBank: AB179623.1).

All previously sequenced variants were tagged with the designation of one reference strain in which they have been sequenced [e.g., the particular variant of SCCmec III from the strain TW20 is indicated SCC [mec III+Cd/Hg+ccrC] (TW20)]. If we were not able to identify a reference sequence to a given SCC hybridization pattern, we added "unknown" followed by the clonal complex(es) and, if there were several similar such elements, by chronologically assigned numbers [as in SCC [mec III+Cd+ccrC] (Unknown, ST239−3)].

Sequencing

The genomes of 30 CC239 isolates from Perth/Australia have been sequenced with Illumina MiSEQ. Sequencing libraries were prepared with the Nextera kit (Illumina).

Genome Assembly

Sequencing reads of 30 CC239 isolates from Perth/Australia as well as several read sets downloaded from NCBI Short read Archive (https://www.ncbi.nlm.nih.gov/sra/; see **Table 4** and **Supplemental Table 1**) were assembled with SPAdes version 3.10.1 (Bankevich et al., 2012). No attempts were made to close gaps between contigs. Contigs shorter than 500 nt were excluded from further analysis.

Bioinformatics, Virtual Hybridizations, and Probe Mapping

To date, several thousand either partially or fully assembled genomes of S. aureus isolates are available in NCBI GenBank. Fully assembled genomes comprise one or several sequences representing complete replicons (the bacterial chromosome and a variable number of plasmids). Partially assembled sequences

consist of a set of contigs. The contigs usually end in repeats and the sequencing reads do not comprise enough information to link contigs unambiguously. The number of contigs varies between about 10 and several hundred depending on the sequencing method, read length, fragment size, coverage depth and assembling strategy and settings. Partially assembled contigs are available in NCBI Genbank (http://www.ncbi.nlm.nih.gov/Traces/wgs/) with special accession numbers assigned which start with four letters followed by eight digits. The entire set of contigs is referred to by a accession number which has all digits set to zero (e.g., AICH00000000.1). For the sake of conciseness, we will refer to these four-letter codes as an unambiguous identifier of a specific genome here. To genomes which we have assembled from raw sequencing reads obtained from the Short Read Archive (https://www.ncbi.nlm.nih.gov/sra/), we refer to henceforth by the BioSample accession number (e.g., SAMEA1029552).

A total of 215 genome sequences of CC239 available in the NCBI database (**Table 4** and **Supplemental Table 1**) as well as genome sequences of 30 previously unpublished Australian study isolates were subjected to an *in silico* analysis or "virtual hybridization" that allowed a direct comparison to array hybridization experiments (**Tables 2a,b, 3** and **Supplemental Table 1**). Hybridization patterns were generated from complete or from partially assembled genomic sequences.

Probe sequences were mapped on contigs using the program blastn (Camacho et al., 2009) from the NCBI blast+ suite and all sites were identified that matched the probe sequences with less than four mismatches. A signal value between 0 and 1 was assigned to each probe based on the actual number of mismatches derived from, and mimicking the normalized signals from a real hybridization experiment. A probe without mismatches was assigned signal intensity of 0.9; with 1 mismatch, a signal of 0.6; with two mismatches, a signal of 0.3; with three mismatches, a signal of 0.1. Probes with four—or more—mismatches were set as 0. These numerical values were then analyzed exactly as data from real hybridization experiments (Monecke et al., 2008b, 2011).

This approach has been developed and optimized based on real experiments performed with fully assembled strains (such as MSSA476, GenBank BX571857; N315, GenBank BA000018; COL, GenBank CP000046; MRSA252, GenBank BX571856; see also Monecke et al., 2016). For three strains (ATCC33592, UK-EMRSA-4, isolate Russia_0085) full genome sequences were available and we have done real as well as virtual hybridization experiments that were analyzed in parallel.

Bioinformatics, Analysis of Insertions

Some CC239 genomes comprise a site-specific insertion of a mobile element into the chromosomale genes *nsaB* (locus tag SATW20_27600) or *yeeE* (locus tag SAT0131_RS10920). Two query sequences with a size of 80 nt were used to evaluate assembled genomes for the presents of uninterrupted *nasB* and *yeeE* genes. The two query sequences were choosen to span the insertion sites (for *nasB*, FN433596.1[2933542:2933621:r] and for *yeeE*, CP002643.1[2151962:2152041:r]). These query sequences were mapped on all full genome sequences with blastn. If they

did not match for their full length, we assumed that the target gene was interrupted.

RESULTS

Subtypes of SCC*mec* III in CC239-MRSA-III

Thirty-nine different variants of SCC*mec* III or SCC*mec* III-derived composite SCC elements or pseudoSCC-elements were observed in the 425 CC239-MRSA-III isolates and sequences. A description of these variants is provided in **Tables 2a,b**. Full profiles for individual isolates and sequences are shown in **Supplemental Table 1**.

Alleles of *mecA* in CC239-MRSA-III were assigned to two alleles matching the CC8-MRSA-I strain COL, CP000046 (among study isolates tested or sequenced, n = 21) and the CC5-MRSA-II strain N315, BA000018 (among study isolates, n = 38; one sequence not unambiguously assigned). When analyzing the binding sites of the probes used, the difference between the alleles is an "A" or, respectively, a "G" in position 737 (of the TW20 *mecA* sequence). Among the sequences and isolates investigated, *mecA* alleles largely correlate with the SCC*mec* subtypes and strains within CC239, i.e., a single *mecA* allele was found in association with each SCC*mec* subtype. However, there were five exceptions, i.e., SCC*mec* subtypes in which both *mecA* alleles were detected. This included the more common subtypes and strains (SCC [*mec* III+Cd/Hg+ccrC] $_{(TW20)}$, SCC [*mec* III+Cd/Hg+ccrC] $_{(Bmb9393)}$, SCC [*mec* III+Cd] $_{(S2)}$, SCC [*mec* III+Cd+*ccrC*] $_{(XN108)}$, SCC [*mec* III+Cd] $_{(HSA10/ATCC33592)}$). This might suggest that the *mecA* sequence, and its allele assignment, is not a reliable phylogenetic marker but subject to random mutation (or to sequencing errors).

All SCC*mec* III elements from CC239 include a cadmium resistance operon for which *cadD* $_{(R35)}$ was used as a marker.

Fifteen SCC*mec* III elements in CC239 ("Eurasian" strains with *ccrC* being located elsewhere not included) were composite elements that additionally harbor the recombinase gene *ccrC*. Sequence analysis identified two different *ccrC* alleles (**Tables 1, 2a,b**) that could not be differentiated with the current set of probes. Genes accompanying *ccrC* are *ccrAA* (although the present allele yields usually only ambiguous signals with the probes used herein) and D1GU38. Nineteen SCC*mec* III elements in CC239 also include the mercury resistance operon; this is often but not always linked to *ccrC*.

Composite elements that include ACME II (that is, *arc* genes present but *opp* genes absent), an arsenic resistance operon, genes *speG* and *czrC* (zinc/cadmium resistance) were occasionally found (see **Tables 2a,b**). No isolates were identified that harbored composite elements involving ACME I (*arc* and *opp* genes), ACME III (*opp* genes only) or SCC*fus* (fusidic acid resistance, *fusC*).

The presence of *dcs* and a SCC terminus sequence or multiple SCC terminus sequences suggests the presence of composite elements. Sequence analyses has shown that SCC terminus sequences are not necessarily situated "terminally" toward *orfX* but can be found within a composite element demarking it components. For example in TW20, SCCterm02 (GenBank FN433596.1; positions 34,140 to 34,456) is situated toward *orfX*

Table 2a | SCCmec-associated patterns observed in, or derived from sequences of, CC239-MRSA assigned to the "Eurasian" (**EA**), "European" (**EU**), "Australian/NZ" (**AU/NZ**) Clades, to the "unassigned Middle Eastern Strain" (**ME**), or to the atypical cluster of "South-East Asian Clade" isolates (S2, DEN907; **SEA**).

SCCmec subtype	Reference sequence	mecA allele	mec complex	Heavy metal resistance genes	Other payload	Recombinase genes	SCC termini	nsaB status	Clade
SCC [mec III+Cd+ccrC] (T0131)	T0131, CP002643.1	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, Q933A2	ccrA+B-3, ccr(AA)+C(PM1)	dcs, SCCterm02	Insertion	EA
SCC [mec III+As/Cd+ccrC] (Unknown, ST239−1)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35), arsC-SCC	D1GU38, Q933A2	ccrA+B-3, ccr(AA)+C	dcs, SCCterm02	N/A	EA
SCC [mec III+Cd+ccrC] (CN79)	CN79, ANCJ	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, D1GU55, Q933A2	ccrA+B-3, ccr(AA)+C(PM1)	dcs, SCCterm02	Insertion	EA
SCC [mec III+Cd+ccrC] [IU17]	IU17, CTXB	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, D1GU55, Q933A2	ccrA+B-3, ccr(AA)+C(PM1)	SCCterm02	Insertion	EA
SCC [mec III+As/Cd+ccrC] (Unknown, ST239−2)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35), arsC-SCC	D1GU38, D1GU55, Q933A2	ccrA+B-3, ccr(AA)+C	dcs, SCCterm02	N/A	EA
SCC [mec III+Cd+ccrC] (Unknown, ST239−1)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2, Q93IB7	ccrA+B-3, ccr(AA)+C	SCCterm01	N/A	ME
SCC [mec III+Cd+ccrC] (Unknown, ST239−2)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, D1GU55, Q4LAG7, Q933A2, Q9S0M4	ccrA+B-3, ccr(AA)+C	dcs, SCCterm02, SCCterm05	N/A	EA
SCC [mec III+Cd+ccrC] (Unknown, ST239−3)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU55, Q4LAG7, Q933A2, Q9S0M4	ccrA+B-3, ccr(AA)+C	dcs, SCCterm05	N/A	EA
SCC [mec III+ ACME II+Cd+ccrC] (Unknown, ST239)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, D1GU55, Q933A2, Q9S0M4, arc-genes	ccrA+B-3, ccr(AA)+C	dcs, SCCterm02, SCCterm03, SCCterm05	N/A	EA
SCC [mec III+Cd+ccrC] [TUR1] (TUR1)	TUR1, SAMEA 1029552	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, D1GU55, Q933A2, Q93IB7	ccrA+B-3, ccr(AA)+C(PM1)	SCCterm01, SCCterm02	Insertion	EA

(Continued)

Table 2a | Continued

SCCmec subtype	Reference sequence	mecA allele	mec complex	Heavy metal resistance genes	Other payload	Recombinase genes	SCC termini	nsaB status	Clade
SCC [mec III+Cd/Hg+ccrC] (SK1585)	SK1585, KL662257.1	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	D1GU38, D1GU55, Q933A2, Q93IB7	ccrA+B-3, ccr(AA)+C(PM1)	SCCterm01, SCCterm02	Wild type	EU
SCC [mec III (Δ mecR1-)+Cd/Hg+ccrC] (BK2421)	BK2421, SAMEA 1029510	mecA(CP000046)	mecA, ugpQ, mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1	mer-operon, cadD(R35)	D1GU38, D1GU55, Q933A2, Q93IB7	ccrA+B-3, ccr(AA)+C(PM1)	SCCterm01, SCCterm02	Wild type	EU
SCC [mec III (Δ mecR1-)+Cd/Hg+ccrC] (85/2082)	85/2082, AB037671.1	mecA(CP000046)	mecA, ugpQ, mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	D1GU38, D1GU55, Q933A2, Q93IB7	ccrA+B-3, ccr(AA)+C(PM1)	SCCterm01, SCCterm02	Wild type	EU
SCC [mec III+Cd] (JKD6008)	JKD6008, CP002120.1	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2	ccrA+B-3	dcs	Wild type	AU/NZ
SCC [mec III+ ACME II+Cd] (Unknown, ST239−1)	N/A	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2, Q9S0M4, arc-genes	ccrA+B-3	dcs, SCCterm03, SCCterm05	N/A	AU/NZ
SCC [mec III+Cd] (S2)	S2, SAMEA 1029563	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1	cadD(R35)	Q933A2	ccrA+B-3	dcs	N/A	SEA
	M116, GenBank CTXQ	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1	cadD(R35)	Q933A2	ccrA+B-3	dcs	N/A	SEA
SCC [mec III+Cd/Hg] (DEN907)	DEN907, SAMEA 1029548	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1	mer-operon, cadD(R35)	Q933A2	ccrA+B-3	dcs	Wild type	SEA
SCCmec III (KM1381), for comparison	KM1381, AM904732.1	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	N/A	Q933A2	ccrA+B-3	dcs	N/A	*pseudint.*

The predicted pattern for S. pseudintermedius, KM1381 is provided for comparison. Bold font of element designations indicates those elements that were found in study isolates.

Table 2b | SCC*mec*-associated patterns observed in, or derived from sequences of, CC239-MRSA assigned to the "South-East Asian" (**SEA**), "South American/Middle Eastern" (**SA/ME**) and "Portuguese" (**POR**) clades.

SCC*mec* subtype	Reference sequence	*mecA* allele	*mec* complex	Heavy metal resistance genes	Other payload	Recombinase genes	SCC termini	*nsaB* status	Clade
SCC [*mec* III+Cd/Hg+ccrC] (TW20)	TW20, FN433596.1	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2, Q93IB7	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm01, SCCterm02	Wild type	SEA, SA/ME
	NMR05, LWAL	*mecA*(BA000018)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2, Q93IB7	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm01, SCCterm02	Wild type	SEA, SA/ME
SCC [*mec* III+Hg+ccrC] (AR23)	N/A	*mecA*(CP000046)	*mecA, ugpQ, mecR1, ΔmecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon	D1GU38, Q933A2, Q93IB7	*ccrA+B-3, ccr(AA)+C*	SCCterm01, SCCterm02	N/A	SA/ME
SCC [*mec* III+Hg] (Unknown, ST239)	N/A	*mecA*(CP000046)	*mecA, ugpQ, mecR1, ΔΔ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon	D1GU38, Q933A2, Q93IB7	*ccrA+B-3*	SCCterm01, SCCterm02	N/A	SEA
SCC [*mec* III+Cd/Hg] (Unknown, ST239−1)	N/A	*mecA*(CP000046)	*mecA, ugpQ, mecR1, ΔmecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2, Q93IB7	*ccrA+B-3*	SCCterm01, SCCterm02	N/A	SEA
SCC [*mec* III+Cd+ccrC] (S85)	S85, SAMEA 1029529	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*cadD*(R35)	D1GU38, Q933A2, Q93IB7	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm01, SCCterm02	Wild type	SEA, SA/ME
SCC [*mec* III+Cd/Hg+ccrC] (Bmb9393)	Bmb9393, CP005288.1	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm02	Wild type	SEA, SA/ME
	URU110, SAMEA 1029540	*mecA*(BA000018)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm02	Wild type	SA/ME
SCC [*mec* III+Cd+ccrC] (PPUKM−775−2009)	PPUKM-775-2009, AMRE	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1,	*cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm02	Wild type	SEA
SCC [*mec* III+Cd/Hg] (Unknown, ST239−2)	N/A	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3*	SCCterm02	N/A	SEA
SCC [*mec* III+Cd/Hg] (BRA2)	BRA2, SAMEA 1029514	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	Q933A2	*ccrA+B-3*	SCCterm02	Wild type	SA/ME
SCC [*mec* III+Cd/Hg+ccrC] (Z172)	Z172, CP006838.1	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1, *mvaS*-SCC	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3, ccr(AA)+C*(TSGH17)	SCCterm02	Wild type	SEA
SCC [*mec* III+Cd/Hg+ccrC] (Unknown ST239)	N/A	*mecA*(CP000046)	*mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB*-SCC1	*mer*-operon, *cadD*(R35)	D1GU38, Q933A2	*ccrA+B-3, ccr(AA)+C*	SCCterm02	N/A	SEA

(Continued)

Table 2b | Continued

SCCmec subtype	Reference sequence	mecA allele	mec complex	Heavy metal resistance genes	Other payload	Recombinase genes	SCC termini	nsaB status	Clade
SCC [mec III+Cd+ccrC] (XN108)	XN108, CP007447.1	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, Q933A2	ccrA+B-3, ccr(AA)+C(TSGH17)	SCCterm02	Wild type	SEA, SA/ME
	DS_014, FQPU	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, Q933A2	ccrA+B-3, ccr(AA)+C(TSGH17)	SCCterm02	Wild type	SA/ME
SCC [mec III+Cd] (HSA10/ATCC33592)	ATCC33592, JXZH	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2, Q93IB7	ccrA+B-3	SCCterm01	Wild type	POR
	P32, SAMEA 862589	mecA(BA000018)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2, Q93IB7	ccrA+B-3	SCCterm01	Wild type	SEA
SCC [mec III+Cd/Hg+ccrC] (M418/UK−9)	M418, CTXS	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	D1GU38, Q933A2, Q9S0M4	ccrA+B-3, ccr(AA)+C	SCCterm02, SCCterm05	Wild type	SEA, SA/ME
SCC [mec III+Cd+ccrC] (NMR08)	CP023560.1	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	D1GU38, Q933A2, Q9S0M4	ccrA+B-3, ccr(AA)+C(TSGH17)	SCCterm02, SCCterm05	Wild type	SEA
SCC [mec III+Cd/Hg] (Unknown, ST239−3)	N/A	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	Q933A2, Q9S0M4	ccrA+B-3	SCCterm05	N/A	SA/ME
SCC [mec III+speG+Cd/czrC+ccrAB4C] (Unknown, ST239)	N/A	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	czrC, cadD(R35)	D1GU38, Q933A2, speG	ccrA+B-3, ccr(AA)+C, ccr(A)/B-4,	SCCterm02, SCCterm07	N/A	SEA
SCC [mec III+speG+Cd/Hg/czrC+ccrAB4C] (Unknown, ST239)	N/A	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, czrC, cadD(R35)	D1GU38, Q933A2, speG	ccrA+B-3, ccr(AA)/C, ccr(A)/B-4	SCCterm02, SCCterm07	N/A	SEA
SCC [mec III+ ACME II+Cd/Hg] (Unknown, ST239)	N/A	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	Q933A2, Q93IB7, Q9S0M4, arc-genes	ccrA+B-3	SCCterm05	N/A	SEA
SCC [mec III+ ACME II+Cd] (Unknown, ST239−2)	N/A	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	cadD(R35)	Q933A2, Q93IB7, arc-genes	ccrA+B-3	SCCterm01, SCCterm03	N/A	SEA
PseudoSCCmec [class A+Cd/Hg+ccrC] (UCIM6015)	UCIM6015, JBMO	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	D1GU38, Q933A2	ccr(AA)+C(TSGH17)	SCCterm01, SCCterm02	Wild type	SEA
PseudoSCCmec [class A+Cd/Hg+ccrC] (UP1073)	UP1073, MIHO	mecA(CP000046)	mecA, ugpQ, mecR1, Δ mecR1, mecI, PSM-mec, xylR/mecR2, cstB-SCC1, mvaS-SCC	mer-operon, cadD(R35)	–	–	SCCterm02	Wild type	SA/ME

Bold font of element designations indicates those elements that were found in study isolates.

(positions 33,660 to 34,139) while SCCterm01 can be found between the region including the mercury resistance operon and the SCCmec element (positions 67,191 to 67,511). As described below, SCCterm02 can also be associated with integration of a transposon into the *nsaB* gene, *i.e.,* at a distant position from *orfX*.

Strains and Clades of CC239-MRSA-III

Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012) divided CC239-MRSA-III into several clades based on SNP analysis. We re-analyzed published sequences from this project, other previously published sequences as well as our own sequences and array hybridization patterns regarding presence of SCCmec III subtypes (see above), the gene *sasX/sesI* (see below), the enterotoxin genes *sek* and *seq* (representative for *S. aureus* pathogenicity island 3), certain resistance markers and regarding other conspicuous features such as *spa* types or deletions of individual genes.

Individual results as well as clade/strain assignments for each isolate and sequence are listed in the **Supplemental Table 1**. A list of strains and clades as well as of their genotypic features is provided in **Table 3**. Strain definitions are mainly based on SCCmec subtypes plus other notable features.

The "Eurasian Clade" and the Insertion into the *nsaB* Gene

Harris and Castillo-Ramírez describe a distinct and rather homogenous "Turkish Clade" (Harris et al., 2010; Castillo-Ramírez et al., 2012). We found that this clade was not restricted to Turkey. It also included isolates and sequences (including T0131, CN79, 16K, 3HK and others) that originate from Eastern Europe and the Balkans, Russia, Pakistan as well as from China (mainly Northern China, but also Hong Kong). Thus we suggest renaming this clade "Eurasian Clade."

Isolates and sequences assigned to this clade are characterized by the presence of $mecA_{(BA000018)}$ and *dcs*. Harris' strains TUR1 and TUR9 differ in this regard being an intermediate to the "European Clade" or a product of a horizontal gene transfer of another variant of a SCCmec III element.

Furthermore, "Eurasian Clade" strains are characterized by integration of a IS431 based transposon—carrying several genes that appear to origin from the SCCmec III element—into the *nsaB* gene which is located 140,000 nt away from the SCC element (T0131, GenBank CP002643: position 2,779,896 to 2,803,726). This transposon consists of *ccrC, ccrAA,* SCCterm02 and additional genes such as *erm*(A), *ant9, hsdR2-* WIS, D1GU60, A9UFT0, Q93IA1, A5INT3, Q9KX75, Q0P7G0, Q93IE0, Q3T2M7, Q4LAG3, D2N370, D1GU38, Q2FKL3, transposase genes and IS431 sequences as well as *aacA-aphD* (absent from T0131, but present in 16K; Yamamoto et al., 2012). This disruption of the *nsaB* gene was described first in isolates from Romania (Chen et al., 2010) and Russia (Yamamoto et al., 2012) but it can be detected in all "Eurasian Clade" sequences. A insertion into *nsaB* is also present in TUR1, SAMEA1029552 and TUR9, SAMEA985415 but due to fragmentation of the sequences into a high number of contigs, the gene content of their insertion cannot be reliably determined.

All other sequences and clades of CC239-MRSA-III present with an un-truncated, wildtype *nsaB* gene. This includes, despite their similarity to the "Eurasian Clade", JKD6008 and related strains from Australia/New Zealand (see below). However, *erm*(A) and *ant9* are not restricted to the insertion into *nsaB*. These two genes can frequently be found in CC239-MRSA strains without the *nsaB* insertion, where they are present in different localizations. In the previously published genome sequence Bmb9393, a transposon carrying these two genes disrupts *radC* (SABB_05268) while in TW20, this transposon is present twice, once integrated into *radC* and once co-localized with the SCCmec III element. In JKD6008, there are also two copies of this transposon, one within *radC* (SAA6008_01621) and one disrupting *ywqG* (SAA6008_00825). Therefore the detection of *erm*(A) and *ant9* cannot be used as a surrogate marker for the identification of the "Eurasian Clade" but the insertion into *nsaB* can.

Other features of the "Eurasian Clade" include a predominance of RIDOM *spa* type t030 (with all isolates previously assigned to *spa* type t030 belonging to this clade; Gostev et al., 2017) or t632 (Moscow and Saint Petersburg), the uniform absence of the mercury resistance operon, the adhesion factor gene *fnbB* and the protease gene *splE* while *splA* and *splB* are present.

The "European Clade"

The basal "European Clade" (Harris et al., 2010) consists of isolates and sequences that share the SCC [mec III+Cd/Hg+ccrC] $_{(SK1585)}$ element and variants thereof from which some genes are fully (*mvaS*) or partially (*mecR1*) deleted. This particular SCCmec element is a composite SCCmec III/heavy metal resistance element that was first observed in a strain isolated in Australia as early as in 1973 (see Nimmo et al., 2015 and section Discussions).

The "European Clade" includes a cluster of homogenous sequences and isolates from, or with epidemiological connection to, Greece. These include Harris' Greek sequences (Harris et al., 2010), isolates from Saxony/Germany epidemiologically linked to Greece (see below and Albrecht et al., 2011) and the Greek reference strain from the Harmony collection (Greece 1_3680). It also includes two isolates from Morocco.

Some "European Clade" strains have a characteristic deletion of 166 nt (in 85/2082, GenBank AB037671.1; corresponding the region in TW20 of FN433596.1 [79030 to 79195]) in the *mecR1* gene that results in the paradoxical observation that the probe associated with *mecR1* yields a signal while the one for ΔmecR1 does not. These include genome sequences from Australia and the US (ANS46, LHH1, BK2421) as well as epidemic strains British UK-MRSA-01 and Irish AR01.

The "South-East Asian Clade" and the *sasX/sesI* Gene

The sequences assigned by Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012) to the (South-East) "Asian Clade" all contain the *sasX/sesI* gene which was absent from all sequences not assigned to this clade. Consequently, *sasX/sesI* was used in the present study as an

Table 3 | Clades and Strains and their genotypic characterization, as identified in this study or when analyzing published genome sequences.

Clade	Strain/isolate name, reference strain or description	Reference sequence	PS	SI	SCCmec element	*cadX*	*blaZ+R+I*	*erm*(A)	*erm*(C)	*aacA-aphD*	*aadD*	*aadE*	*ant9*	*aphA3+sat*	*ble*	*dfrA*	*dfrG*	*mupA*	*tet*(K)	*tet*(M)	*cat*	*qacA*	*qacC*	*sek+seq*	*sesX/sesI*	*Hlb*-integrating phage	Other notable features
Eurasian	TUR1	SAMEA1029552	2	0	SCC [*mec* III+Cd+*ccrC*]$_{\text{(TUR1)}}$	○	●	●	●	●	○	○	●	○	○	○	○	○	●	●	◉	○	●	○	○	*sea+sak+scn*	***splE* and *fnbB* absent**
	T0131	CP002643.1	4	1	SCC [*mec* III+Cd+*ccrC*]$_{\text{(T0131)}}$	○	●	●	◉	○	○	○	●	○	○	○	○	○	○	●	○	○	◉	●	○	*sea+sak+scn*	***splE* and *fnbB* absent**
	***arsC*-SCC-pos. strains from Romania**	–	0	2	SCC [*mec* III+As/Cd+*ccrC*]$_{\text{(Unknown, ST239−1)}}$ SCC [*mec* III+As/Cd+*ccrC*]$_{\text{(Unknown, ST239−2)}}$	○	●	●	○	◉	○	○	●	○	○	○	○	○	○	●	○	○	○	●	○	*sea+ sak+ scn*	***splE* and *fnbB* absent**
	16K/CN79	GenBank ANCJ, BABZ	60	28	SCC [*mec* III+Cd+*ccrC*]$_{\text{(CN79)}}$	○	●	⦾	◉	⦾	⊙	○	⦾	●	⊙	⊙	○	⊙	○	●	⊙	○	◉	⦾	○	*sea+sak+scn* (comm.); *sea*$_{\text{(N315)}}$+ *sak* + *chp* + *scn* (rare, Saint Petersburg only); none (rare)	***splE* and *fnbB* absent**
	IU17	GenBank CTXB	2	0	SCC [*mec* III+Cd+*ccrC*]$_{\text{(IU17)}}$	○	●	●	●	●	○	○	●	○	○	○	○	○	○	●	○	○	○	●	○	*sea+sak+scn*	***splE and fnbB absent***
	Sporadic strains from Saxony	–	0	2	SCC [*mec* III+Cd+*ccrC*]$_{\text{(Unknown, ST239−2)}}$ SCC [*mec* III+Cd+*ccrC*]$_{\text{(Unknown, ST239−3)}}$	○	●	●	◉	●	○	○	●	○	○	○	○	○	○	●	○	○	○	●	○	*sea*	***splE* and *fnbB* absent**
	Kuwaiti ACME II-pos. Strain	–	0	2	SCC [*mec* III+ ACME II+Cd+*ccrC*]$_{\text{(Unknown, ST239)}}$	○	●	●	●	●	○	○	◉	○	○	○	○	○	○	●	○	○	○	●	○	*sea*	***splE* and *fnbB* absent**. ACME II present
European	**"Greek Epidemic Strain"**	SAMEA1029513, SAMEA1029532	3	7	SCC [*mec* III+Cd/Hg+*ccrC*]$_{\text{(SK1585)}}$	◉	●	○	⦾	⦾	○	⦾	○	⦾	○	○	○	○	◉	⦾	⊙	○	⊙	⊙	○	*sea+sak+scn*	*fnbA* commonly absent
	Irish AR01/UK EMRSA-01	SAMEA1029537, AB037671.1 (SCC*mec* only)	2	5	SCC [*mec* III (Delta *mecR*1-)+ Cd/Hg+*ccrC*]$_{\text{(85/2082)}}$	○	●	●	○	⦾	○	○	●	○	○	◉	○	○	●	●	◉	◉	○	●	○	*sea+sak+scn*	*bbp* absent from UK1 and Irish isolates
	BK2421	SAMEA1029510	1	0	SCC [*mec* III (Delta *mecR*1-)+ Cd/Hg+*ccrC*]$_{\text{(BK2421)}}$	○	●	●	○	●	○	○	●	●	○	●	○	○	●	●	○	●	○	●	○	*sea+sak+scn*	-
Australian/NZ	**JKD6008**	CP002120.1	2	29	SCC [*mec* III+Cd]$_{\text{(JKD6008)}}$	○	●	●	○	⦾	⦾	⦾	●	○	●	◉	○	○	○	●	○	◉	○	◉	○	*sea+sak+scn* (var.); or none	*splA+B+E* variable
	ACME-pos. variant of JKD6008	–	0	3	SCC [*mec* III+ ACME II+Cd]$_{\text{(Unknown,ST239−1)}}$	○	●	●	○	●	●	○	●	○	●	●	○	○	○	●	○	●	◉	○	○	*sea+sak+scn*	ACME II present
South-East Asian	**TW20 (Irish AR44)**	GenBank FN433596.1	37	36	SCC [*mec* III+Cd/Hg+*ccrC*]$_{\text{(TW20)}}$	◉	⦾	●	⊙	⦾	⊙	●	⦾	●	○	⊙	◉	⊙	◉	●	⊙	◉	◉	⦾	●	*sea+sak+scn* (comm.); *sak+chp+scn* (var.); *sak+scn* (rare) none (rare)	*splE* absent from Mideastern isolates and Syrian sequence M592, SAMEA862550

(Continued)

Table 3 | Continued

Gene/marker columns (left to right): cadX, blaZ+R+I, erm(A), erm(C), aacA-aphD, aadD, aadE, ant9, aphA3+sat, ble, dfrA, dfrG, mupA, tet(K), tet(M), cat, qacA, qacC, sak+seq, sasX/sesI

Clade	Strain/isolate name, reference strain or description	Reference sequence	PS	SI	SCCmec element	Hlb-integrating phage	Other notable features
	Strain with a TW20-derived pseudoelement	GenBank JBMO	1	1	PseudoSCCmec [class A+Cd/Hg+ccrC] (UCIM6015)	sea+sak+scn or none	–
	Z172	GenBank CP006838.1	1	1	SCC [mec III+Cd/Hg+ccrC] (Z172)	sea+sak+scn (var); or none	splE variable
	mvaS-deletion variant of Z1721		0	1	SCC [mec III+Cd/Hg+ccrC] (UnknownST239)	sea+sak+scn	–
	XN108	GenBank CP007447.1	4	9	SCC [mec III+Cd+ccrC] (XN108)	sea+sak+scn (var); sak+chp+scn (rare); none (rare)	–
	S85	SAMEA1029529	2	3	SCC [mec III+Cd+ccrC] (S85)	sak+chp+scn	–
	PPUKM-775-2009-	GenBank AMRE	1	1	SCC [mec III+Cd+ccrC] (PPUKM-775-2009)	sea+sak+scn or sak+chp+scn	–
	sasX-pos. mer+/ccrC neg. isolates from Kuwait	–	0	2	SCC [mec III+Hg] (Unknown, ST239)	sea+sak+scn	–
	sasX-pos. strain with a Bmb9393-like SCC element	GenBank CP013957.1 (Bmb9393)	18	17	SCC [mec III+Cd/Hg+ccrC]	sea+sak+scn (var); sak+chp+scn (rare); sak+scn (rare) none (rare)	–
	M418	GenBank CTXS	1	0	SCC [mec III+Cd/Hg+ccrC] (M418/UK–9)	sea+sak+scn	–
	NMR08	GenBank CP023560.1	2	1	SCC [mec III+Cd+ccrC] (NMR08)	sea+sak+scn	–
	sasX-pos. ACME-pos. strain	–	0	1	SCC [mec III+ ACME II+Cd/Hg] (Unknown, ST239)	sea+sak+scn	ACME II present
	sasX-pos/czrC-pos. strains	–	0	5	SCC [mec III+speG+Cd/Hg/czrC +ccrAB4C] (Unknown, ST239) SCC [mec III+speG+Cd/czrC+ccrAB4C] (Unknown, ST239)	none	–

(Continued)

Table 3 | Continued

Clade	Strain/isolate name, reference strain or description	Reference sequence	PS	SI	SCCmec element	cadX	blaZ+R+I	erm(A)	erm(C)	aacA-aphD	aadD	aadE	ant9	aphA3+sat	ble	dfrA	dfrG	mupA	tet(K)	tet(M)	cat	qacA	qacC	sek+seq	sesX/sesI	Hlb-integrating phage	Other notable features
	sasX-pos. sporadic strains from Kuwait and Hong Kong		0	5	SCC [*mec* III+Cd/Hg] (Unknown, ST239−1) SCC [*mec* III+Cd/Hg] (Unknown, ST239−2)	◉ ○	●	●	○	○	○	●	●	●	●	○	○	○	○	●	○	○ ◉	○	●	● ○	*sea*+*sak*+*chp*+*ssl01* to *ssl05*; *sea*+*sak*+*scn* or none	HK isolate
	P32	GenBank CTXV	1	0	SCC [*mec* III+Cd] [HSA10/ATCC33592]	○	●	●	○	●	○	●	●	●	●	●	●	○	●	●	○	○	○	●	●	none	*sdrD* absent
	ACME-pos. strain similar to P32	GenBank CTXV	1	1	SCC [*mec* III+ACME II+Cd] (Unknown, ST239−2)	○	●	●	○	●	○	●	●	●	●	●	●	○	●	●	○	○	○	●	●	*sak*+*scn*	–
hla-negative South-East Asian	DEN907	SAMEA1029548	1	0	SCC [*mec* III+Cd/Hg] [DEN907]	●	●	●	○	●	○	●	●	●	●	○	○	●	○	●	○	○	○	●	●	*sea*+*sak*+*scn*	**hla absent**
	S2/M116	SAMEA1029563 GenBank CTXQ	12	0	SCC [*mec* III+Cd] [S2]	◐	○	●	○	●	◉	◉	●	◉	○	○	○	●	○	●	○	◉	●	●	●	*sea*+*sak*+*scn*	**hla absent**
Portuguese	**Portuguese & ATCC 33592 strain**	GenBank JXZH	8	5	SCC [*mec* III+Cd] [HSA10/ATCC33592]	○	○	◉	◉	◉	◉	◉	●	◉	◉	◉	○	○	○	●	◉	◉	◉	◉	○	*sea*+*sak*+*scn* (rare); *sak*+*chp*+*scn* (rare); *sak*+*scn* (var) or none (var)	*spl* genes variably absent, *bbp* and *sdrD* sporadically absent
South American/ Middle Eastern	**Bmb9393/UK EMRSA-11**	GenBank CP005288.1	34	4	SCC [*mec* III+Cd/Hg+ccrC] (Bmb9393)	◉	●	●	○	◉	○	◉	●	●	○	◉	◉	●	○	●	◉	◉	●	◉	○	*sak*+*chp*+*scn* (comm.) or none	
	BRA2-Strain	SAMEA1029514	6	0	SCC [*mec* III+Cd/Hg] [BRA2]	○	●	●	◉	◉	○	●	●	●	○	○	○	●	○	●	●	○	○	○	○	*sak*+*chp*+*scn* (comm.) or none	
	Strain with a BRA2-derived pseudoelement	GenBank MIHO	1	0	PseudoSCC*mec* [class A+Cd/Hg] (UP1073)	○	○	●	○	●	○	○	○	●	○	○	○	●	●	●	○	○	○	○	○	*sak*+*chp*+*scn*	
	MRSA_PR1	GenBank ANPO	2	3	SCC [*mec* III+Cd+ccrC] [XN108]	◉	●	●	○	◉	○	○	○	○	○	◉	●	○	○	●	◉	○	○	●	○	*sea*+*sak*+*scn* (var.); *sak*+*chp*+*scn* (var.)	*bbp* absent in MRSA_PR1 but present in others
	sasX-negative strain with a S85-like SCC element	–	0	2	SCC [*mec* III+Cd+*ccrC*] [S85]	●	●	●	○	●	○	○	○	○	○	○	●	○	●	●	○	○	○	●	○	*sak*+*chp*+*scn*	*splE* absent
	UK EMRSA-09	–	0	1	SCC [*mec* III+Cd/Hg+ccrC] (UK−EMRSA−9)	○	●	●	○	●	○	○	○	○	○	○	●	○	●	●	○	○	○	○	○	*sak*+*scn*	

(Continued)

Table 3 | Continued

Clade	Strain/isolate name, reference strain or description	Reference sequence	PS	SI	SCCmec element	Hlb-integrating phage	Other notable features
	Isolate from Togo	–	0	1	SCC [mec III+Cd/Hg] (Unknown, ST239–3)	sea+sak+scn	
	Irish AR09/ UK EMRSA-04 and –07	GenBank AEEK	7	20	SCC [mec III+Cd/Hg+ccrC] (TW20)	sea+sak+scn (var); sak+chp+scn (var.) sea(N315)+sak+chp+ scn (rare, Algeria only); none (rare)	splE variably absent, ssl01 to ssl06 absent in HK isolates
Related to South American/ Middle Eastern	Irish AR23	–	0	1	SCC [mec III+Hg+ccrC] (AR23)	sea+sak+scn	
	Krasnoyarsk tst1 + Strain	GenBank BBKC	1	4	SCC [mec III+Cd/Hg+ccrC] (Bmb9393) with mecA[BA000018]	sak+scn	tst1 positive
	URU110	SAMEA1029540	1	0	SCC [mec III+Cd/Hg+ccrC] (Bmb9393) with mecA[BA000018]	sea+sak+scn	
	DS_014	GenBank FQPU	1	0	SCC [mec III+Cd+ccrC] (XN108) with mecA[BA000018]	sak+chp+scn	
Unassigned	Unassigned Middle Eastern Strain	–	0	4	SCC [mec III+Cd+ccrC] (Unknown, ST239–1)	sea+sak+scn	hla absent

Gene columns (variable genes): cadX, blaZ+R+I, erm(A), erm(C), aacA-aphD, aadD, aadE, ant9, aphA3+sat, ble, dfrA, dfrG, mupA, tet(K), tet(M), cat, qacA, qacC, sek+seq, sesX/sesI.

Bold font of element designations indicates those elements that were found in study isolates. Column "PS" indicates the number of (non-duplicate) previously published sequences analyzed, the Column "SI" indicates the number of study Isolates. Symbols/abbreviations for frequencies of variable genes: ○, always absent; ⊙, rare, i.e., present in <20% of analyzed sequences and/or characterized isolates; ◉, variable, i.e., present in 20–80%; ●, common, i.e., present in more than 80%; ●, always present.

identifying marker for this clade. In all *sasX/sesI*-positive CC239 sequences analyzed, the *sasX/sesI* prophage is localized in the same position within the genome, splitting *yeeE* (FN433596.1; positions 2180899 to 2181684 downstream of the phage insertion, and 2308889 to 2309177 upstream). This, and the observation that the entire cluster appeared in previous sequencing studies to be monophyletic (Harris et al., 2010; Castillo-Ramírez et al., 2012) suggest that the "South-East Asian Clade" is one distinct lineage resulting from one single acquisition of the *sasX/sesI* prophage. Differences affecting mobile genetic elements (including SCC*mec*) and the presence of the alpha haemolysin gene *hla* (which is absent from several sequences, mainly from the Middle East and Thailand) could then be considered secondary.

SCC*mec* elements in the "South-East Asian Clade" are complex composite elements consisting of SCC*mec* III [usually, but not always, with *mecA*(CP000046)], *ccrC*, *cadD* as well as of SCCterm01 and/or 02 sequences. The mercury resistance operon is nearly always present, and the few exceptions might be regarded as secondary deletions.

Isolates and sequences originate from South-East Asia, including Hong Kong and Southern China, India, Australia, the Middle East, Western Europe and, surprisingly, Trinidad & Tobago. The clade comprises also TW20 (NCTC13626), AUS-EMRSA-3 and Harmony Collection Finland E24_98541.

The "South American/Middle Eastern Clade"

Furthermore, there is a large clade encompassing a wide variety of isolates and sequences and that show identical or very similar SCC*mec* types as the "South-East Asian Clade" (**Table 2b**) but that are *sasX/sesI*-negative. This includes the "South American Clade" sequences (Harris et al., 2010; Castillo-Ramírez et al., 2012) with three different SCC*mec* subtypes (SCC [*mec* III+Cd/Hg+*ccrC*] (Bmb9393), SCC [*mec* III+Cd/Hg] (BRA2) without *ccrC* and associated genes and pseudoSCC*mec* [class A+Cd/Hg] (UP1073) without any *ccr* genes). However, there are also isolates and sequences that originate mainly from the Middle East, but also from Europe and Russia, and that have identical or similar SCC*mec* types. Hence, we referred to this clade herein as the "South American/Middle Eastern Clade." This clade includes Bmb9393, ATCC BAA-39, NCTC13131, UK-EMRSA-4, UK-EMRSA-7, UK-EMRSA-9 and UK-EMRSA-11, Irish AR09 and AR23 and the unique *tst1*-positive strain from Krasnoyarsk, Russia.

Other Clades

Furthermore, there are additional geographically restricted clades and strains that do not fit into the larger clades as defined by Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012).

One group of sequences and isolates could be named the "Australian/New Zealand Clade" consisting of a number of isolates from Australia and New Zealand with JKD6009 and JKD6008 being representative genome sequences. Isolates lacked *sasX/sesI*, the *mer* operon and *ccrC*. They carried *mecA*(BA000018). Similar to the "Eurasian Clade" strains, they harbored *dcs* but they differed in the absence of the secondary SCC-like gene cluster inserted into *nsaB*.

Another clade comprises a very homogenous cluster of Portuguese genome sequences from Harris' work and a similar strain, ATCC 33592, from New York City. The isolates and sequences lack *sasX/sesI, dcs, ccrC*, the mercury resistance operon and an integration into *nsaB*, but harbor SCCterm01.

We also observed a cluster of isolates from the Middle East, Libya and Russia that did not match any published sequences. These isolates carried *mecA*(BA000018) and *ccrC* but lacked the mercury resistance operon, the *nsaB* integration, *sasX/sesI* and *hla*.

Isolates by Geographic Origin

An overview on geographic origins by clade and strain is provided in **Table 4**. In the following paragraph a short summary by countries and different sampling sites is given.

African Countries

Although CC239-MRSA-III has been observed in several African contries (Jansen van Rensburg et al., 2011; Abdulgader et al., 2015), there are insufficient data on epidemiological trends or molecular epidemiology.

Six isolates from five different African countries were included in the study. One isolate from Algeria was a "South American/Middle Eastern Clade" strain carrying SCC*mec* III+Cd/Hg+*ccrC* (TW20), and was identical to Irish AR09. One isolate from a Libyan patient (who was brought to North-Eastern Germany for humanitarian aid) belonged to the unassigned "Middle Eastern" strain. Two isolates from Morocco matched the "European Clade", being most similar to the "Greek Strain" but differed from the other isolates in the presence of *sek/seq* and the absence of *cadX, erm*(C) and *tet*(M). One isolate from Togo matched the "South American/Middle Eastern Clade" but harbored an unique SCC*mec* subtype [designated SCC*mec* III+Cd/Hg (Unknown, ST239−3) in **Table 2b**]. This isolate might be a derivative of UK EMRSA-9 that has lost *ccrC* and accompanying genes. One isolate from Uganda (Monecke et al., 2013a) was *sasX/sesI*-positive and belonged to the "South-East Asian Clade."

Australia

CC239-MRSA-III have been present in Australia for decades (Dubin et al., 1991; Coombs et al., 2004; Howden et al., 2010; Nimmo et al., 2015) with distinct variants (Aus-2 EMRSA and Aus-3 EMRSA) being distinguished based on mercury susceptibility or resistance, respectively (Coombs et al., 2006).

Eighteen Australian isolates were genotyped by microarray, and Ilumina NGS sequences of an additional 30 isolates were analyzed. The majority of isolates ($n = 31$) were very similar to JKD6008, GenBank CP002120.1, and JKD6009, GenBank ABSA from New Zealand and Australia, respectively, forming a distinct "Australian/New Zealand (NZ) Clade" (correlating to the mercury susceptible "Aus-2 EMRSA"). Three isolates harbored an ACME II cluster as well as Q9S0M4 and yielded additional signals for SCCterm03 and 05. The isolates might be regarded as direct derivatives of the JKD6008/6009 strain.

Thirteen isolates belonged to the "South-East Asian Clade" possibly indicating foreign importation which, given the geographic links between Australia and Asia, seems likely.

Table 4 | Clades and strains of CC239-MRSA-III, geographic origin of isolates and sequences, GenBank and BioSample accession numbers.

Clade	Strain/Isolate	Origin of study isolates	Origin of previously published sequences	Fully assembled genome sequences (GenBank)	WGS (Whole Genome Shotgun) genomes (Prefixes)	BioSample accession numbers of self-assembled genomes
Eurasian	TUR1	–	Turkey	–	–	SAMEA1029552,–985415
	T0131	Romania (1)	Turkey, China,	CP002643.1	CTYE, CTYI, CTYV	–
	arsC-SCC-pos. strains from Romania	Romania (2)	–	–	–	–
	16K/CN79	Germany (from Turkey; 1), Germany (from Makedonia; 1), Germany (1), Romania (7); Russia (Moscow; 5); Russia (Saint Pietersburg; 4); Russia (Chelyabinsk; 2); Russia (Kurgan; 2); Pakistan (5), China (Hong Kong; 2)	Czech Republic, Hungary, Romania, Russia, Turkey, Pakistan, China	–	ANCJ, BABZ, AZMV, AZMX, CTWB, CTWC, CTWD, CTWE, CTWF, CTWG, CTWH, CTWI, CTWJ, CTWK, CTWL, CTWM, CTWN, CTWO, CTWP, CTWQ, CTWT, CTWX, CTXD, CTXF, CTXJ, CTXK, CTYB, CTYF, CTYH, CTYJ, CTYK, CTYL, CTYM, CTYN, CTYO, CTYP, CTYR, CTYS, CTYT, CTYU, CTYW, CTYX, CTYY, CTYZ, CTZC, CTZF, CTZG, CTZI, CTZJ, CTZK, CTZM, CTZO, CTZP, JTJX	SAMEA1029551, –1029557, –1029566, –1029567
	IU17		Turkey	–	CTXB, CTXE	–
	Sporadic strains from Saxony	Germany (from Middle East; 1), Germany (1)	–	–	–	–
	Kuwaiti ACME II-pos. Strain	Kuwait (2)	–	–	–	–
European	"Greek Epidemic Strain"	Germany (from Greece, 4), Greece (Harmony collection; 1), Morocco (2)	Greece, USA (R35)	KL662257.1 (SCCmec only)		SAMEA1029513, –1029532, –1029546, –1029564
	UK-01/Irish AR01 Epidemic Strain	Ireland (4), UK (NCTC11939; 1)	Australia, USA	AB037671.1 (SCCmec only)		SAMEA1029537,–1029523
	BK2421	–	USA	–		SAMEA1029510
Australian/NZ	JKD6008-Strain	Australia (31)	Australia, New Zealand	CP002120.1	ABSA	–
	ACME-pos. variant of JKD6008-Strain	Australia (3)	–	–	–	–
South-East Asian	TW20 (Irish AR44)	Ireland (2), Kuwait (3), Saudi Arabia (3), China (Hong Kong; 9), Australia (6), Uganda (1), Trinidad&Tobago (13)	UK, Germany, Denmark, Syria, Sri Lanka, India, Thailand, Malaysia, China (Nanjing), China (Hong Kong), USA	FN433596.1, CP015447.1	AZJQ, CTWW, CTWZ, CTXH, CTXM, CTXN, CTXT, CTXU, CTXZ, CTYD, FFPZ, FGZX, FQJD, FQJP, JBEK, JBGY, JBNV, JBOE, JBPU, JBRH, JBSL, JVZJ, JVZK, JVZL, LWAL	SAMEA1029512, –1029516, –1029519, –1029524, –1029528, –1029538, –1029541, –1029560, SAMN04903357

(Continued)

Table 4 | Continued

Clade	Strain/Isolate	Origin of study isolates	Origin of previously published sequences	Fully assembled genome sequences (GenBank)	WGS (Whole Genome Shotgun) genomes (Prefixes)	BioSample accession numbers of self-assembled genomes
	Strain with a TW20-derived pseudoelement	China (Hong Kong; 1)	N/A	–	JBMO	–
	Z172	Russia (Kurgan, 1)	Taiwan	CP006838.1	–	–
	mvaS-deletion variant of Z1721	Australia (1)	–	–	–	–
	XN108	Germany (from India; 1), Saudi Arabia (1), India (3), China (Hong Kong; 3), Trinidad&Tobago (1)	India, China (Chongqing), Thailand, Malaysia	CP007447.1	FQFW, JQNU, LWBY	–
	S85	Saudi Arabia (3)	Thailand	–	CTWU	SAMEA1029529
	PPUKM-775-2009	China (Hong Kong; 1)	Malaysia	–	AMRE	–
	sasX-pos. mer+/ccrC neg. isolates from Kuwait	Kuwait (2)	–	–	–	–
	sasX-pos. strain with a Bmb9393-like SCC element (AUS-EMRSA-3)	Finland (Harmony collection; 1), Kuwait (1), Saudi Arabia (1), India (1), Pakistan (1), China (Hong Kong; 7), Australia (5), Trinidad&Tobago (1)	Lebanon, India, Hong Kong, Korea, Thailand, Malaysia	CP013957.1	AHZL, AICJ, AMRB, AMRC, AVPR, AZMZ, CTWV, CTXL, CVRW, FGBR, FQJI, FQJW, LWAK, MLQB, NESW	SAMEA1029556, –1029568
	M418	–	India	–	CTXS	–
	NMR08	India (1), China (Hong Kong; 1)	India	CP023560.1	LWAM	–
	sasX-pos. ACME-pos. strain	Australia (1)	–	–	–	–
	sasX-pos/czrC-pos. strains	Trinidad&Tobago (5)	–	–	–	–
	sasX-pos. sporadic strains from Kuwait and Hongkong	Kuwait (4), China (Hong Kong; 1)	–	–	–	–
	P32	–	Poland	–	CTXV	–
	ACME-pos. strain similar to P32	Australia (1)	–	–	–	–
hla-negative South-East Asian	DEN907	–	Denmark	–	–	SAMEA1029548

(Continued)

Table 4 | Continued

Clade	Strain/Isolate	Origin of study isolates	Origin of previously published sequences	Fully assembled genome sequences (GenBank)	WGS (Whole Genome Shotgun) genomes (Prefixes)	BioSample accession numbers of self-assembled genomes
	S2/M116	–	Thailand, Vietnam	–	CTXQ	SAMEA1029511, −1029515, −1029522, −1029525, −1029535, −1029536, −1029547, −1029553, −1029558, −1029563, −1029565
Portuguese	**Portuguese & ATCC 33592 strain**	Russia (Chelyabinsk;1), Russia (Kurgan; 3) USA (ATCC 33592)	Portugal, USA	–	JXZH	SAMEA1029520, −1029533, −1029543, −1029545, −1029561, −1029562, −1029542
South American/ Middle Eastern	**Bmb9393/UK EMRSA-11**	UK (UK EMRSA-11; 1), Germany (from Middle East; 1), Ecuador (2)	UK, Germany, Denmark, France, Spain, Portugal, Czech Republic, Lithuania, Brazil, Uruguay, Argentina	CP005288.1, CP009681.1, CP012011.1, CP012012.1, CP012013.1, CP012015.1, CP012018.1	AHVC, AICH, CTXC, CTXG, CTXI, CTXO, CTXP, CTXW, CTXX, CTXY, CTYA, CTYC, CTYG, LDPC, LGWS	SAMEA1029508, −1029509, −1029517, −1029527, −1029534, −1029544, −1029549, −1029550, −1029554, −1029559, −1029569, −985416
	BRA2-Strain	–	Brazil, Chile, Peru	–	LGWM, LGWY, LGWZ	SAMEA1029514, −1029526, −1029539
	Strain with a BRA2-derived pseudoelement	–	Peru	–	MIHO	–
	MRSA_PR1	Saudi Arabia (3)	Malaysia, Lithuania	–	ANPO, CTXR	–
	sasX-negative strain with a S85-like SCC element	Saudi Arabia (2)	–	–	–	–
	UK EMRSA-09	UK (UK EMRSA-09)	–	–	–	–
	Isolate from Togo	Togo (1)	–	–	–	–
	Irish AR09/ UK EMRSA-04 and−07	UK (UK EMRSA-04 and−07), Ireland (3), Russia (Krasnoyarsk; 1), Saudi Arabia (9), Kuwait (1), Algeria (1), China (Hong Kong; 1), Australia (1)	UK, Denmark, Hungary, Egypt		AEEK, CTWR, CTWS, FFVG	SAMEA1029518, −1029531, −3529263
	Irish AR23	Ireland (1)	–	–	–	–
Related to South American/ Middle Eastern	**Krasnoyarsk tst1+ Strain**	Russia (Krasnoyarsk; 4)	Russia	–	BBKC	–
Unassigned	URU110	–	Uruguay	–	FQPU	SAMEA1029540
	DS_014	–	Thailand	–	–	–
	Unassigned Middle Eastern Strain	Germany (from Libya, 1), Saudi Arabia (1), Kuwait (1), Russia (Moscow; 1)	–	–	–	–

Bold font indicates that isolates belonging to those strains have been found in the study.

Interestingly, two of these isolates also harbored the ACME II cluster. One isolate matched the Middle Eastern/Irish AR09 strain. One previously published Australian CC239 genome sequence, ANS46 (SAMEA1029537) belonged to the "European Clade," UK-01/AR01.

The SCC*mec* element of the "Greek Strain," SCC [*mec* III+Cd/Hg+*ccrC*] $_{(SK1585)}$, was also previously observed in Australia, in the chimeric strain, ST2249-MRSA-III, SK1585 (GenBank AYLT and KL662257.1) (Nimmo et al., 2015).

China/Hong Kong

CC239-MRSA-III has been epidemic in China for decades. In Hong Kong, a presence of the "South-East Asian Clade" was reported while in Northern China, "Eurasian Clade" strains emerge and spread (Ip et al., 2005; Chen et al., 2010, 2014a; Wang et al., 2014).

Twenty-seven isolates originated from Hong Kong. The majority ($n = 23$) belonged to the "South-East Asian Clade" with the most common strains carrying SCC [*mec* III+Cd/Hg+*ccrC*] $_{(TW20)}$ or SCC [*mec* III+Cd/Hg+*ccrC*] $_{(Bmb9393)}$ (nine and seven isolates, respectively). Another isolate had a SCC [*mec* III+Cd/Hg+*ccrC*] $_{(TW20)}$-derived pseudoSCC*mec* element. Two isolates belonged to the "Eurasian Clade", likely to indicate influx from mainland China (Wang et al., 2014) and two isolates were assigned to the "South American/Middle Eastern Clade."

Ecuador

In Ecuador, 2005-2013, CC239-MRSA-III was the second most common MRSA strain (Zurita et al., 2016) and contrarily to other Latin American countries it is still common there (Arias et al., 2017).

Two isolates were assigned to the "South American/Middle Eastern Clade" being *sasX/sesI* negative while carrying SCC [*mec* III+Cd/Hg+*ccrC*] $_{(Bmb9393)}$.

Germany/Saxony

CC239-MRSA-III is not an epidemic strain in the German state of Saxony, or at least it has not been since 2000 and many cases are related to foreign travel (Albrecht et al., 2011).

Eleven isolates were included into the study. Nine of them were obtained from patients with known travel history (including admission to foreign healthcare facilities), with nosocomial contact to travelers, or from immigrant patients.

Five isolates were assigned to the "Eurasian Clade." Two of them were obtained from Macedonian and Turkish nationals, respectively, the latter with history of hospitalization in Turkey after trauma. A third patient with an "Eurasian Clade" isolate appeared to have a Middle Eastern background while for two remaining cases, no history of immigration or travel was known. The "Greek Strain" CC239-MRSA-[III+Cd/Hg+*ccrC*] $_{(SK1585)}$ was found in four outbreak isolates, with an index patient who was repatriated from Greece after trauma and emergency care (Albrecht et al., 2011). One patient with a "South American/Middle Eastern Clade" strain had a Middle Eastern background indeed. Finally, one isolate from a patient of Indian background belonged to the "South-East Asian Clade" and matched the SCC*mec* element of XN108.

India

CC239-MRSA-III appear to be common and widespread in India; and although other strains emerged meanwhile, it is, at least regionally, still a dominant MRSA strain (D'Souza et al., 2010; Abimanyu et al., 2012; Neetu and Murugan, 2016).

In addition to one isolate from an Indian patient in Saxony (see above), five isolates with an Indian background were tested. All belonged to the "South-East Asian Clade"; three matched genome sequence of XN108 and its SCC*mec* subtype, and one the Indian genome sequences of NMR07/08. A fifth isolate was *sasX/sesI* positive but had a Bmb-9393-like SCC*mec* element.

Ireland

CC239-MRSA-III predominated in Irish hospitals in the mid-to-late 1980s [locally known as phenotype III and antibiogram-resistogram (AR) types 01 and 09] but has since only been recovered sporadically or as part of localized outbreaks, represented by AR15 and AR23 isolates recovered in 1992/93 and AR44 recovered in 2002 (Carroll et al., 1989; Rossney et al., 1994; Shore et al., 2005). The 10 Irish ST239-MRSA-III isolates investigated clustered into three clades and four strains.

Firstly, AR01/AR15 isolates with a distinct truncation of *mecR1* matched Harris' "Basal/European Clade," being identical to sequences of Ans46 and LHH1 (from the US and Australia) as well as to UK-EMRSA-1.

Secondly, AR09/Phenotype III isolates harbored SCC [*mec* III+Cd/Hg+*ccrC*] $_{(TW20)}$, lacked *sasX/sesI* and matched Middle Eastern isolates. This fits to the observation that this strain was first brought to Ireland with an oil worker who was repatriated from Iraq in 1985 with a subsequent major outbreak (Humphreys et al., 1990).

AR23 could be considered a variant of the AR09 strain that was *cadD*-negative.

Thirdly, AR44 harbored SCC [*mec* III+Cd/Hg+*ccrC*] $_{(TW20)}$ and were *sasX/sesI*-positive. It has been suggested previously that this strain was imported from Singapore (Rossney, 2003), and indeed that this variant predominates in South-East Asia. This outbreak was contained and did not spread beyond one unit (Shore et al., 2005).

Kuwait

In Kuwait, CC239-MRSA-III accounted for more than 50% of typed MRSA isolates collected in a period from 1992 to 2010 (Boswihi et al., 2016). Currently, it is still present although it appears to be replaced by community-acquired MRSA strains (Udo and Al-Sweih, 2017).

Fourteen Kuwaiti isolates belonged to nine different strains and were assigned to the "Eurasian," "South American/Middle Eastern," and the "South-East Asian" clades. Two isolates of the "Eurasian Clade" harbored ACME II elements thus differing from all other isolates of that clade. One was identical to the Irish AR09 outbreak strain that was reported to originate from the Middle East (see above and Humphreys et al., 1990). Three "South-East Asian Clade" isolates were essentially identical to TW20 and one had a Bmb9393-like SCC*mec* element. Six others represented sporadic variants that were characterized by a loss of *ccrC* although the usually accompanying D1GU38 was present.

One isolate belonged to an unassigned strain that was found mainly in the Middle East.

Pakistan

There are few studies on genotyping of MRSA from Pakistan indicating a presence of CC239-MRSA-III in hospitals (Shabir et al., 2010; Zafar et al., 2011; Arfat, 2013; Jamil et al., 2017) but its absence in the community.

Five Pakistani isolates from Rawalpindi (as well as the one previously published genome sequence from Pakistan, NCTR #32S, GenBank JTJX) belonged to the "Eurasian Clade", CN79/16K-Strain. One isolate was assigned to "South-East Asian Clade" strain with a BMB939-like SCC*mec* element.

Romania

CC239-MRSA-III matching the "Eurasian Clade" has been reported to be common in Romania after 2000 (Cirlan et al., 2005; Chen et al., 2010; Monecke et al., 2014a).

All 10 Romanian isolates included, as well as previously published sequences, clustered into the "Eurasian Clade", lacking *sasX/sesI* but carrying *mecA*(BA000018), *dcs* as well as SCCterm02 (indicating the secondary SCC-like gene cluster inserted into *nsaB*). Seven matched genome sequences 16K and CN79, and one T0131. Two isolates had unsequenced composite SCC*mec* elements including the arsenic resistance gene *arsC* that might be regarded as variants of 16K/CN79- and T0131-like elements, respectively.

Russia

After 2000, CC239-MRSA-III was found to be common in different regions across Russia, not only in hospitals but also in the community (Afanas'ev et al., 2010; Baranovich et al., 2010; Yamamoto et al., 2012; Khokhlova et al., 2015; Gostev et al., 2017). Previous reports (Afanas'ev et al., 2010; Baranovich et al., 2010; Yamamoto et al., 2012; Gostev et al., 2017) and sequences (16 K) indicated a presence of the "Eurasian Clade" as well as of other variants (Gostev et al., 2017). Besides, there was a notable emergence of a distinct *tst1*-positive variant in the city of Krasnoyarsk (Khokhlova et al., 2015).

The majority of Russian isolates (13/24), from Moscow, Saint Petersburg, Kurgan and Chelyabinsk, matched the "Eurasian Clade" and genome sequences CN79 and 16 K. However, three isolates from Saint Petersburg differed from the others in the presence of the *sea*(N315)/*sep* allele. Four isolates from Krasnoyarsk represent a local epidemic strain harboring *tst1* and SCC [*mec* III+Cd/Hg+*ccrC*] (Bmb9393). Isolates were identical (although one isolate lacked presumably plasmid-borne *cat*, encoding chloramphenicol resistance) to the genome sequence of MRSA-OC3, GenBank BBKC, SAMD00019145 which also originated from this town. The four isolates from Kurgan and Chelyabinsk were essentially identical to ATCC 33592 (representing a clade previously known from Portugal and the USA). One isolate from Kurgan was identical to the Taiwanese genome sequence Z172 (SAMN02370325). One isolate from Krasnoyarsk belonged to the "South American/Middle Eastern Clade" and one isolate from Moscow matched the unassigned strain that was otherwise found in the Middle East.

Saudi Arabia

Although CC239-MRSA-III is known to be present in the Middle Eastern/Gulf region for decades (Humphreys et al., 1990), molecular data confirming a presence of CC239-MRSA-III in the Kingdom of Saudi Arabia have been published only in recent years (Cirlan et al., 2005; Al-Obeid et al., 2010; Monecke et al., 2012c; Senok et al., 2016), and differences in carriage of *ccrC*, *merA/B* and aminoglycoside resistance genes indicated a simultaneous existence of different variants of this strain (Monecke et al., 2012c).

Twenty-three isolates from two different hospitals in Riyadh were characterized. Fourteen were assigned to the "South American/Middle Eastern Clade", nine had SCC [*mec* III+Cd/Hg+*ccrC*] (TW20) thus matching Irish AR09 (see above and Humphreys et al., 1990). However, the 14 contemporary Saudi Arabian isolates lacked *splE*. As isolates were obtained from two hospitals in one city this might indicate a recent outbreak situation. The differences in SCC*mec* subtypes (absence of the *mer* operon resulting in SCC [*mec* III+Cd+*ccrC*] (S85) and of *mer*, SCCterm01 and Q93IB7 resulting in SCC [*mec* III+Cd+*ccrC*] (XN108)) would then only be secondary to the loss of *splE*.

Another eight isolates belonged to the "South-East Asian Clade" (being also all *splE*-negative) and one belonged to the unassigned "Middle Eastern Strain."

Trinidad & Tobago

CC239-MRSA-III has been the dominant MRSA clone in Trinidad & Tobago at least from 2000/01 (Akpaka et al., 2007) until 2012/2013 (Monecke et al., 2012b, 2014b).

Twenty-one isolates from Trinidad & Tobago where characterized. All were *sasX/sesI*-positive. This observation places CC239 from Trinidad & Tobago into the "South-East Asian Clade," rather than any other Latin American strains. The majority ($n = 14$) carried SCC [*mec* III+Cd/Hg+*ccrC*] (TW20). Four isolates harbored composite SCC elements that included *czrC*, *speG* and additional recombinase genes *ccrA/B*-4. Based on the overall patterns, the isolates could be regarded as derivatives of SCC [*mec* III+Cd/Hg+*ccrC*] (Bmb9393).

DISCUSSION

Evolution of the SCC*mec* III Element

CC239 is closely linked to SCC*mec* III. Only a few other lineages carry this type of SCC*mec* element. These include CC5 (Kwon et al., 2006; Shittu et al., 2009; Monecke et al., 2013b), CC398 (Nemati et al., 2008) and *Staphylococcus pseudintermedius* (KM1381, GenBank AM904732.1; KM241, GenBank AM904731). Virtually all CC239-MRSA isolates harbor one of the many different subtypes and composite elements of SCC*mec* III.

A relatively simple SCC*mec* III element, i.e., harboring *ccrAB3* and a class A *mec* complex without additional *ccr* genes, heavy metal resistance markers, integrated transposons etc. has only been observed in *S. pseudintermedius* (KM1381, GenBank AM904732.1). However, it cannot safely be assumed that SCC*mec* III was initially transmitted from

S. intermedius/pseudintermedius as, to the best of our knowledge, the earliest observation of methicillin resistance in "*S. intermedius*" was reported in 1984 (Roy et al., 1984). The most similar SCC*mec* III element in *S. aureus* can be found in the Sanger sequenced "Eurasian Clade" strain T0131 (CP002643) where it is only supplemented by the integration of a cadmium resistance operon (for which *cadD*(R35) is used as marker herein). The secondary set of SCC markers—*ccrC, ccrAA*, SCCterm02, D1GU38, *erm*(A) and *ant9*—are integrated elsewhere in the genomes of this lineage, distant from SCC*mec* and *orfX*. This is not only an interesting oddity, but raises the very practical question whether other SCC and SCC*mec* elements exist at alternative chromosomal sites away from *orfX*. If they do, this would have major consequences for rapid molecular MRSA tests as these assays target the integration of SCC*mec* into *orfX*.

Another relatively simple SCC*mec* element can be found in JKD6008 (CP002120.1) that, however, harbors additional resistance genes *cadD*(R35), *aadD* and *ble*. Hybridization patterns consistent with this element were observed in most "Australian/NZ Clade" isolates and in a livestock-associated CC5-MRSA-III isolate (see Monecke et al., 2013b and **Supplemental Table 1**). KM1381, JKD6008 as well as all "Australian/NZ Clade" and "Eurasian Clade" isolates and sequences (except TUR1 and TUR9; Harris et al., 2010) have *dcs* rather than other SCC terminal sequences and *mecA*(BA000018).

TUR1/TUR9 and strains of the "European" (85/2082, AB037671.1), "South-East Asian" (TW20, FN433596.1; XN108, CP007447; Z172, CP006838.1) and "South American/Middle Eastern" (Bmb9393, CP005288.1) clades harbor more complex SCC*mec* elements that include the cadmium resistance operon, *ccrC, ccrAA* and D1GU38 as well as (often, but not always) *erm*(A) and *ant9, tet*(K) and the mercury resistance operon. These isolates do not have *dcs* and a vast majority of isolates carry *mecA*(CP000046). "European Clade" sequences share the *ccrC* allele (*ccrC* (PM1)) with the "Eurasian Clade" while "South-East Asian" and "South American/Middle Eastern" clades have, if present, a different *ccrC* allele [*ccrC* (TSGH17)].

The presence of the *mer* genes raises the question for the benefit of mercury resistance in *S. aureus*/MRSA. One possible explanation could be the past medical use of mercury (e.g., for the treatment of syphilis, in topical agents such as merbromin, or in dental restorative materials such as amalgam) that could pose a selective pressure also on staphylococci colonizing the patients in question, regardless of whether they belonged to *S. aureus* or to other, coagulase-negative staphylococcal species. This could mean that SCC*mer* elements predated SCC*mec* in the same way as mercury use predated the clinical use of antibiotics. If SCC*mec* elements evolved indeed already after the introduction of penicillin (Harkins et al., 2017), there may have been a couple of decades of time for the evolution and selection of composite SCC*mec*/SCC*mer* elements.

The composite SCC [*mec* III+Cd/Hg+*ccrC*] (SK1585) element (SK1585, KL662257.1) existed at least already in the very early 1970s (as it was found in a strain epidemic in Australia from 1973 on; see below) and it appears to be ancestral to many SCC elements in CC239-MRSA. All SCC elements in "European," "South-East Asian," and "South American/Middle

Eastern" clades could easily be described as variants of this particular element that have either acquired additional genes (ACME II, *speG, czrC, ccrA/B4*) or lost some or several genes. These latter genes have either no known function, are redundant (in case of the transposon with *erm*(A) and *ant9* as a second copy is present elsewhere in the genome) or may no longer be of major advantage anymore because the compounds that provide a selective pressure are no longer frequently used (as in the case of mercury).

When mapping the presence of subtypes of SCC*mec* III on the phylogenetic trees as proposed by Harris and Castillo-Ramírez (Harris et al., 2010; Castillo-Ramírez et al., 2012), it becomes clear that identical subtypes can be observed in different clades (e.g., SCC [*mec* III+Cd/Hg+*ccrC*] (TW20), SCC [*mec* III+Cd/Hg+*ccrC*] (Bmb939)). There might be two different, mutually non-exclusive explanations. Firstly, these elements are subject to horizontal transfer so that SCC*mec* elements may be lost, acquired and exchanged after differentiation into different clades. Secondly, many of the SCC*mec* III subtypes differ in losses or acquisitions of accessory, purposeless or redundant genes (see above), and such events may have occurred several times. For instance, SCC [*mec* III+Cd/Hg+*ccrC*] (TW20) and SCC [*mec* III+Cd/Hg+*ccrC*] (Bmb9393) differ only in absence of the redundant transposon carrying *erm*(A) and *ant9,* Q93IB7 and SCCterm01 from the latter. It seems to be possible that this loss (or other similar losses) may have happened multiple times, independently from each other, to different lineages harboring SCC [*mec* III+Cd/Hg+*ccrC*] (TW20). Among a cluster of "South American/Middle Eastern Clade" isolates from Riyadh with a characteristic *splE* deletion, we observed three SCC*mec* subtypes suggesting that the losses of the mercury operon, SCCterm01 and Q93IB7 are secondary only to the loss of this gene and that they may spontaneously change SCC [*mec* III+Cd/Hg+*ccrC*] (TW20) to SCC [*mec* III+Cd+*ccrC*] (S85) and SCC [*mec* III+Cd+*ccrC*] (XN108).

An example for multiple acquisitions of one gene cluster is the observation of ACME II in dissimilar SCC elements of rather unrelated Australian and Kuwaiti strains. Likewise, repeated and independent acquisitions of *arc* genes have already been observed in Singapore (Hsu et al., 2015).

Evolution and Spread of the CC239-MRSA-III Strain

Based on accumulation of SNPs and mutation rates, previous work (Harris et al., 2010; Castillo-Ramírez et al., 2012) estimated the emergence of CC239 to have occurred in the mid- or late 1960s. The preservation of CC239-MRSA isolated in 1971 in a Norwegian strain collection (Smyth et al., 2010) also hints to an emergence and early spread of this strain in the late 1960s to 1970.

When analyzing gene content, one needs to assume two major recombination events to have occurred. One event was a horizontal gene transfer of a large segment of CC30 DNA into a CC8 genome (Robinson and Enright, 2004b; Holden et al., 2009). The other was a transfer of a SCC*mec* III element either before that "CC8/CC30 hybridization" into the CC30 ancestral strain, or afterwards, into the CC239 chimeric strain. It is not clear

which gene transfer happened first. We are not aware of a CC30-MRSA-III strain that may have posed as a donor for the CC30 core genomic DNA and the SCC*mec* III element. CC239-MSSA strains have been identified (Strain 21178, GenBank AGRN and Luedicke et al., 2010), but they might be secondary deletion variants that lost SCC*mec* III rather than methicillin-susceptible ancestors to CC239-MRSA-III.

Our own observations also indicate that at least the "Greek Strain" and its SCC [*mec* III+Cd/Hg+*ccrC*] (SK1585) must have existed already in the early 1970s. There are reports of CC239 from Australia from this time and another strain, ST2249-MRSA-III was present in Melbourne, Australia, from 1973-1979 (predating the oldest Australian isolates of CC239 by 3 years). ST2249-MRSA-III is a chimeric strain (Nimmo et al., 2015) that combined features of CC45, CC30, and CC8 parental strains. The CC30- and CC8-like parts of its genome can be seen as one continuous segment originating from a CC239 parental strain, also including the SCC [*mec* III+Cd/Hg+*ccrC*] (SK1585) element that is characteristic for the "Greek Strain". This allows two assumptions. Firstly, an importation of the "Greek Strain" of CC239 (or of the ST2249 chimera after the hybridization event) from Greece to Melbourne appears not improbable given a large Greek community in this city. Secondly, if the recombination that gave raise to ST2249-MRSA-III happened in 1973 or earlier, the "Greek Strain" CC239 and its SCC*mec* III element must have existed already some time before allowing for its emergence and spread as far as Australia. As discussed above, the SCC [*mec* III+Cd/Hg+*ccrC*] (SK1585) element could conveniently be regarded ancestral to many SCC*mec* elements in "European," "South-East Asian," and "South American/Middle Eastern" clades assuming that these elements emerged by serial or multiple deletions, and, occasionally, by acquisitions of genes. The other "European Clade" strains characterized by a distinct *mecR1* deletion may have evolved in the early 1980s and spread in a rather limited way, i.e., in Ireland, UK, Australia, New Zealand and the USA during that decade (Ito et al., 2001; Shore et al., 2005; Harris et al., 2010). "European Clade" strains have been in recent years still of some relevance in Greece (and in travelers returning from there), but otherwise they have been replaced by other MRSA strains.

Then there is the "Eurasian Clade" (or Harris' and Castillo-Ramírez' "Turkish Clade"; Harris et al., 2010; Castillo-Ramírez et al., 2012). A comparatively low number of distinct strains within this clade might indicate a rather recent emergence, and earliest sequences identified (Harris et al., 2010) originate from Eastern Europe and Turkey, from the mid-/late 1990s. Genotyping data indicate relatedness to the "European Clade" (Harris et al., 2010; Castillo-Ramírez et al., 2012) but the "Eurasian Clade" and the "Australian/NZ Clade" differ from others by harboring less complex composite SCC elements with *dcs* and *mecA*(BA000018). Whether this indicates an independent, second acquisition of SCC*mec* III cannot yet be determined. The absence of *splE* and *fnbB* from all isolates and sequences indicate a monophyletic, clonal origin of the entire clade. TUR1 and TUR9 differ from other strains suggesting yet another horizontal gene transfer (possibly of a SCC*mec* III element from a "European" strain into an "Eurasian" strain with interrupted *nsaB*). The

"Eurasian Clade" can be found in Turkey where it is frequently isolated and widespread (Tekeli et al., 2016). Furthermore it occurs in Eastern Europe including Macedonia (from where one study patient came from, see paragraph on Saxony/Germany) and, especially, Romania. In Hungary, it was common in the 1990s but it is declining since then, being replaced by other strains (Conceicao et al., 2007). It is also present in Russia, Pakistan and China. Recent reports from China indicate an emergence of the "Eurasian Clade" (with *spa* t030) at the expense of other CC239 strains (that is, of the "South-East Asian Clade") following a North South gradient. Its distribution within China suggests import from Central Asia and/or a spill-over across the Russian border (Chen et al., 2014b). It appears to replicate faster than the "South-East Asian Clade" strains (Shang et al., 2016) and this advantage appears to outweigh in direct competition whatever advantage the presence of *sasX/sesI* may confer to the "South-East Asian" strains.

As mentioned, we found isolates matching Harris' and Castillo-Ramírez' "South American Clade" also in Russia and the Middle East. This raises the question where it emerged and to where it spread secondarily. Castillo-Ramírez estimated "*the introduction into South America to have occurred approximately…in 1992 (late 1989, 1993)*" (Castillo-Ramírez et al., 2012). Since the Irish AR09/Phenotype III outbreak strain was brought to Ireland from Iraq in 1985, and since it was described to be similar to a strain sampled in Baghdad as early as 1984 (Humphreys et al., 1990) we assume that this clade evolved earlier, possibly in the Middle East from where it may have spread to India, Russia and Europe. Again, travel from India to the Middle East and back as well as from the Middle East to Europe might have played a role. Strains of this clade may have come to Latin America from Europe or directly from the Middle East, and it became common and widespread in several Latin American countries (Harris et al., 2010; Castillo-Ramírez et al., 2012). Recent evidence, however, shows that this clade is declining or disappearing, except possibly in Ecuador and Peru (Arias et al., 2017). Many sequences of the "South American/Middle Eastern Clade" originated from Brazil (Harris et al., 2010; Castillo-Ramírez et al., 2012). While it is tempting to assume a link between Portugal and Brazil, a majority of Portuguese sequences clearly belong to a separate, geographically restricted, clade; and the few sequences that match the "South American/Middle Eastern Clade" might be re-imported by travelers (Harris et al., 2010; Castillo-Ramírez et al., 2012). While it is receding in Latin America and in India (D'Souza et al., 2010) and while it largely disappeared from Ireland, this clade still appears to be endemic in the Middle East and in Russia. One notable strain carrying *tst1* has been endemic in the Russian town of Krasnoyarsk for several years (first observed in 2008; Iwao et al., 2012). The presence of the *tst1* gene in CC239 is rather unique although this or a similar strain has also been described from Iran (Havaei et al., 2013).

The "South-East Asian Clade" most likely evolved from a "South American/Middle Eastern Clade" strain (or from a common ancestor of both clades) by acquisition of a prophage carrying *sasX/sesI*. Providing that this gene was acquired only once, which is the most parsimonious assumption, it might be assumed that this happened between the split of the related

"South-East Asian," "Portuguese," and "South American/Middle Eastern" lineages and the proliferation of different strains within "South-East Asian Clade", i.e., between ca. 1969 and 1985 based on Castillo-Ramírez' data (Castillo-Ramírez et al., 2012). The oldest published genome sequences originate from 1997 (CUHK_HK1997), 1998 (CHI59), 2001 (DEN907), and 2003 (TW20). The "South-East Asian Clade" spread in South-East Asia, including India, Thailand, Malaysia, Singapore, and China. Although it is still present in Hong Kong and in Southern Mainland China, in Northern China "Eurasian Clade" strains predominate nowadays (see above and Chen et al., 2014b; Shang et al., 2016). The "South-East Asian Clade" was also occasionally introduced to Europe, most likely by travelers (DEN907, TW20, AR44, P32, Finland E24_98541 from the Harmony collection), without becoming endemic there, and it also has been identified in Canada and the USA. The presence of the "South-East Asian Clade" and particularly of strains that appear to originate from India and South-East Asia in Kuwait and Saudi Arabia may easily be attributed to the large number of Indian and South-East Asian workers in the Gulf States (Birks et al., 1988). An interesting observation is the presence of this clade, rather than of the "South American" one, on the Caribbean islands of Trinidad & Tobago. A possible explanation is the Indian/South Asian descent of a high proportion of inhabitants of Trinidad & Tobago (ca. 38% of the total population, or 1.4 million people; http://www.tt.undp.org/content/dam/trinidad_tobago/docs/DemocraticGovernance/Publications/TandT_Demographic_Report_2011.pdf). Another 35% are of African descent, but no sufficient subtyping data for African CC239-MRSA are available. Possibly, an importation of MRSA by visits to ancestral lands might have played a greater role in the case of Trinidad & Tobago than just the mere geographic proximity to Latin America.

Some of Harris' strains from Thailand, one from Vietnam and one from Denmark are placed into the "South-East Asian Clade" by presence of *sasX/sesI* and by sequence analysis (Harris et al., 2010; Castillo-Ramírez et al., 2012). However, they differ from other strains of that clade in lacking *hla* and in harboring *dcs* instead of other SCC terminal sequences. The latter could indicate that their SCC*mec* elements rather originated from a horizontal gene transfer, maybe from the Australian/NZ lineage (see **Table 2a**).

Finally, there are some isolates and sequences that do not fit into the major clades. This includes the "Portuguese Clade" and the "Australian/New Zealand Clade." The former is, according to sequence analysis (Harris et al., 2010; Castillo-Ramírez et al., 2012), related to the "South-East Asian" and "South American/Middle Eastern" clades. The latter was not represented by SNP-based studies and its SCC*mec* element might be more related to the one in non-CC239 strains (including *S. pseudintermedius* KM1381) than to the SCC*mec* elements in other CC239. We identified a cluster of Middle Eastern isolates (including one from Libya and one from Russia) that might constitute yet another clade. Finally there are strains such as DS_014, UR110 and P32 that could be assigned to the major clades but that still differ from them in particular features (such as *mecA* alleles). They may have evolved by further horizontal gene transfers. They also could be representatives of separate lineages or clades of CC239 that may be restricted to certain geographic regions poorly, or not at all, covered by previous typing and sequencing work. It might be expected that there are even more such unrecognized clades because CC239 was common in Western Europe and the USA before modern typing and sequencing technologies emerged, and because it is now common in countries were such technologies are not extensively applied.

Regarding typing technologies, NGS methods and DNA array hybridization profiling allow assignment to clades and strains. Arrays are currently cheaper and more convenient in a clinical setting. NGS can achieve a higher resolution although the definition of a "breakpoint for identity or non-identity" (i.e., how many differences between related isolates rule out direct transmission) still poses a challenge. This is quite a relevant issue for practical purposes. Traditionally, a "group of isolates that can be distinguished from other isolates of the same genus and species by phenotypic characteristics or genotypic characteristics or both" were regarded as a strain or clone (Tenover et al., 1995; Dijkshoorn et al., 2000). However, recent typing technologies achieve a level of resolution that is sufficiently informative to differentiate dozens of variants within one "strain" such as CC239-MRSA-III (as seen in the tables herein). Therefore, defining "strains" may still be useful for epidemiological purposes, but it is somewhat awkward and prone to subjectivity. Both approaches, short read NGS methods and microarray hybridization profiling, have difficulties with gene duplications and translocations if potentially mobile genes are flanked by repetitive and multi-copy sequences. Practically, this means that both technologies are useful for typing, but for the reconstruction of phylogenetic relationships, conventional sequencing still is unsurpassed.

On a very practical level, the definition of clades or variants can be useful for infection control purposes. For this pandemic strain it was possible to define such clades and to link molecular identifiers to geographic origins. Analyses of markers discussed herein, regardless whether by array hybridization, multiplex PCR, or by genome sequencing, can help assigning clinical isolates to these clades or variants and thus help to identify the provenance of an isolate and to discern imported from locally acquired cases. This is relevant as this strain was able to cause large hospital-born outbreaks upon importation with travelers or repatriated patients, as for instance the Irish experience (Humphreys et al., 1990; Shore et al., 2005), the TW20 outbreak in London (Holden et al., 2009), our own observations of the "Greek Strain" in Saxony or the spread of the "South American/Middle Eastern" clade in Latin America or of the "Eurasian Clade" in China showed. Based on European and North American experience, it is tempting to assume that CC239-MRSA-III has been side-lined by other clones or has even become extinct. Given the increasing scale of global travel and migration, there is still a possibility of re-importation and secondary spread. One should keep in mind that this strain still frequently detected in hospitals serving literally more than half of the world's population, i.e., China, India, South-East Asia, Turkey and the Middle East, Romania, Russia and parts of Latin America.

In conclusion, CC239-MRSA-III is a truly pandemic strain that, for nearly half a century, traveled around the world, infecting and even killing thousands of patients. This pandemic does not originate from elusive animals hosts in jungles and savannahs but from professionals working in the cleanest and most hygienic environments possible, that is, hospitals and operating theaters. Typing techniques allow following these movements, and even pinpointing individual index patients from whom this strain was brought into certain countries. However, understanding of a pandemic does not automatically results in an ability to prevent it. The very fact that an exclusively hospital-borne pandemic can spread that far and can last that long emphasizes an urgent need for improved hand hygiene, mandatory screening of staff and admitted patients, and decolonization procedures, a prudent use of antimicrobial agents and in general far more effective infection prevention and control measures.

AUTHOR CONTRIBUTIONS

SM designed the study, supervised and analyzed experiments, and wrote the manuscript. PS designed the primers and probes for the arrays used herein and analyzed genome sequences as well as experimental data. DG, EM, AR, AR-L, and RR performed experiments. SB and VG performed experiments and obtained isolates. PA, DB, MB, OD, MI, BJ, LJ, MN, AS, SS, LS, AMS, MS, AT, EU, TV, and JZ obtained isolates and provided clinical/epidemiological data. DC, GC, and ACS obtained isolates, provided clinical/epidemiological data and revised the manuscript. RE designed the study, supervised experiments, and revised the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01436/full#supplementary-material

Supplementary Table 1 | Full hybridization profiles for isolates characterized by microarray, and predicted hybridization profiles for analyzed sequences.

REFERENCES

Abdulgader, S. M., Shittu, A. O., Nicol, M. P., and Kaba, M. (2015). Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in Africa: a systematic review. *Front. Microbiol.* 6:348. doi: 10.3389/fmicb.2015.00348

Abimanyu, N., Murugesan, S., and Krishnan, P. (2012). Emergence of methicillin-resistant *Staphylococcus aureus* ST239 with high-level mupirocin and inducible clindamycin resistance in a Tertiary Care Center in Chennai, South India. *J. Clin. Microbiol.* 50, 3412–3413. doi: 10.1128/JCM.01663-12

Afanas'ev, M. V., Karakashev, S. V., Il'ina, E. N., Salem Al-Salami, A. M., Sidorenko, S. V., and Govorun, V.M. (2010). [Molecular characteristic of methicillin-resistant *Staphylococcus aureus* strains isolated in Moscow clinics]. *Mol. Gen. Mikrobiol. Virusol.* 2, 20–24.

Aires de Sousa, M., Bartzavali, C., Spiliopoulou, I., Sanches, I. S., Crisóstomo, M. I., and de Lencastre, H. (2003a). Two international methicillin-resistant *Staphylococcus aureus* clones endemic in a university hospital in Patras, Greece. *J. Clin. Microbiol.* 41, 2027–2032. doi: 10.1128/JCM.41.5.2027-2032.2003

Aires de Sousa, M., Crisostomo, M. I., Sanches, I. S., Wu, J. S., Fuzhong, J., Tomasz, A., et al. (2003b). Frequent recovery of a single clonal type of multidrug-resistant *Staphylococcus aureus* from patients in two hospitals in Taiwan and China. *J. Clin. Microbiol.* 41, 159–163. doi: 10.1128/JCM.41.1.159-163.2003

Akpaka, P. E., Kissoon, S., Rutherford, C., Swanston, W. H., and Jayaratne, P. (2007). Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* isolates from regional hospitals in Trinidad and Tobago. *Int. J. Infect. Dis.* 11, 544–548. doi: 10.1016/j.ijid.2007.03.004

Albrecht, N., Jatzwauk, L., Slickers, P., Ehricht, R., and Monecke, S. (2011). Clonal replacement of epidemic methicillin-resistant *Staphylococcus aureus* strains in a German university hospital over a period of eleven years. *PLoS ONE* 6:e28189. doi: 10.1371/journal.pone.0028189

Al-Obeid, S., Haddad, Q., Cherkaoui, A., Schrenzel, J., and François, P. (2010). First detection of an invasive *Staphylococcus aureus* STRAIN (D958) with reduced susceptibility to glycopeptides in Saudi Arabia. *J. Clin. Microbiol.* 48, 2199–2204. doi: 10.1128/JCM.00954-09

Alp, E., Klaassen, C. H., Doganay, M., Altoparlak, U., Aydin, K., Engin, A., et al. (2009). MRSA genotypes in Turkey: persistence over 10 years of a single clone of ST239. *J. Infect.* 58, 433–438. doi: 10.1016/j.jinf.2009.04.006

Arfat, Y. (2013). *Genotyping of Methicillin Resistant Staphylococcus aureus (MRSA) from Local Hospital of Rawalpindi/Islamabad, Pakistan.* Ph.D., Quaid-i-Azam University, Islamabad.

Arias, C. A., Reyes, J., Carvajal, L. P., Rincon, S., Diaz, L., Panesso, D., et al. (2017). A prospective cohort multicenter study of molecular epidemiology and phylogenomics of *Staphylococcus aureus* bacteremia in nine Latin American countries. *Antimicrob. Agents Chemother.* 61:e00816-17. doi: 10.1128/AAC.00816-17

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Baranovich, T., Zaraket, H., Shabana, I. I., Nevzorova, V., Turcutyuicov, V., and Suzuki, H. (2010). Molecular characterization and susceptibility of methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* isolates from hospitals and the community in Vladivostok, Russia. *Clin. Microbiol. Infect.* 16, 575–582. doi: 10.1111/j.1469-0691.2009.02891.x

Becker, K., van Alen, S., Idelevich, E. A., Schleimer, N., Seggewiß, J., Mellmann, A., et al. (2018). Plasmid-encoded transferable *mecB*-mediated methicillin resistance in *Staphylococcus aureus*. *Emerging Infect. Dis.* 24, 242–248. doi: 10.3201/eid2402.171074

Birks, J. S., Seccombe, I. J., and Sinclair, C. A. (1988). Labour migration in the Arab Gulf states: patterns, trends and prospects. *Int. Migr.* 26, 267–286. doi: 10.1111/j.1468-2435.1988.tb00649.x

Boswihi, S. S., Udo, E. E., and Al-Sweih, N. (2016). Shifts in the clonal distribution of methicillin-resistant *Staphylococcus aureus* in Kuwait Hospitals: 1992-2010. *PLoS ONE* 11:e0162744. doi: 10.1371/journal.pone.0162744

Budimir, A., Deurenberg, R. H., Bosnjak, Z., Stobberingh, E. E., Cetkovic, H., and Kalenic, S. (2009). A variant of the Southern German clone of methicillin-resistant *Staphylococcus aureus* is predominant in Croatia. *Clin. Microbiol. Infect.* 16, 1077–1083. doi: 10.1111/j.1469-0691.2009.03042.x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Campanile, F., Bongiorno, D., Borbone, S., and Stefani, S. (2009). Hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) in Italy. *Ann. Clin. Microbiol. Antimicrob.* 8:22. doi: 10.1186/1476-0711-8-22

Carroll, J. D., Pomeroy, H. M., Russell, R. J., Arbuthnott, J. P., Keane, C. T., McCormick, O. M., et al. (1989). A new methicillin- and gentamicin-resistant *Staphylococcus aureus* in Dublin: molecular genetic analysis. *J. Med. Microbiol.* 28, 15–23. doi: 10.1099/00222615-28-1-15

Castillo-Ramírez, S., Corander, J., Marttinen, P., Aldeljawi, M., Hanage, W. P., Westh, H., et al. (2012). Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol.* 13:R126. doi: 10.1186/gb-2012-13-12-r126

Cha, H. Y., Moon, D. C., Choi, C. H., Oh, J. Y., Jeong, Y. S., Lee, Y. C., et al. (2005). Prevalence of the ST239 clone of methicillin-resistant *Staphylococcus aureus* and differences in antimicrobial susceptibilities of ST239 and ST5 clones identified in a Korean Hospital. *J. Clin. Microbiol.* 43, 3610–3614. doi: 10.1128/JCM.43.8.3610-3614.2005

Chen, H., Yang, X., Wang, Q., Zhao, C., Li, H., He, W., et al. (2014a). Insights on evolution of virulence and resistance from the whole-genome analysis of a predominant methicillin-resistant *Staphylococcus aureus* clone sequence type 239 in China. *Chin. Sci. Bull.* 59, 1104–1112. doi: 10.1007/s11434-014-0149-1

Chen, L., Mediavilla, J. R., Smyth, D. S., Chavda, K. D., Ionescu, R., Roberts, R. B., et al. (2010). Identification of a novel transposon (Tn6072) and a truncated staphylococcal cassette chromosome mec element in methicillin-resistant *Staphylococcus aureus* ST239. *Antimicrob. Agents Chemother.* 54, 3347–3354. doi: 10.1128/AAC.00001-10

Chen, Y., Liu, Z., Duo, L., Xiong, J., Gong, Y., Yang, J., et al. (2014b). Characterization of *Staphylococcus aureus* from distinct geographic locations in china: an increasing prevalence of spa-t030 and SCCmec type III. *PLoS ONE* 9:e96255. doi: 10.1371/journal.pone.0096255

Cirlan, M., Saad, M., Coman, G., Bilal, N. E., Elbashier, A. M., Kreft, D., et al. (2005). International spread of major clones of methicillin resistant *Staphylococcus aureus*: nosocomial endemicity of multi locus sequence type 239 in Saudi Arabia and Romania. *Infect. Genet. Evol.* 5, 335–339. doi: 10.1016/j.meegid.2004.09.005

Conceicao, T., Aires-de-Sousa, M., Fuzi, M., Toth, A., Paszti, J., Ungvari, E., et al. (2007). Replacement of methicillin-resistant *Staphylococcus aureus* clones in Hungary over time: a 10-year surveillance study. *Clin. Microbiol. Infect.* 13, 971–979. doi: 10.1111/j.1469-0691.2007.01794.x

Coombs, G. W., Nimmo, G. R., Bell, J. M., Huygens, F., O'Brien, F. G., Malkowski, M. J., et al. (2004). Genetic diversity among community methicillin-resistant *Staphylococcus aureus* strains causing outpatient infections in Australia. *J. Clin. Microbiol.* 42, 4735–4743. doi: 10.1128/JCM.42.10.4735-4743.2004

Coombs, G. W., Pearson, J. C., O'Brien, F. G., Murray, R. J., Grubb, W. B., and Christiansen, K. J. (2006). Methicillin-resistant Staphylococcus aureus clones, Western Australia. *Emerg Infect Dis.* 12, 241–247. doi: 10.3201/eid1202.050454

Cuevas, O., Cercenado, E., Bouza, E., Castellares, C., Trincado, P., Cabrera, R., et al. (2007). Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in Spain: a multicentre prevalence study (2002). *Clin. Microbiol. Infect.* 13, 250–256. doi: 10.1111/j.1469-0691.2006.01672.x

Dijkshoorn, L., Ursing, B. M., and Ursing, J. B. (2000). Strain, clone and species: comments on three basic concepts of bacteriology. *J. Med. Microbiol.* 49, 397–401. doi: 10.1099/0022-1317-49-5-397

D'Souza, N., Rodrigues, C., and Mehta, A. (2010). Molecular characterization of Methicillin resistant *Staphylococcus aureus* (MRSA) with emergence of epidemic clones ST 22 and ST 772, in Mumbai, India. *J. Clin. Microbiol.* 48, 1806–1811. doi: 10.1128/JCM.01867-09

Dubin, D. T., Matthews, P. R., Chikramane, S. G., and Stewart, P. R. (1991). Physical mapping of the mec region of an American methicillin-resistant *Staphylococcus aureus* strain. *Antimicrobial Agents Chemother.* 35, 1661–1665.

Edgeworth, J. D., Yadegarfar, G., Pathak, S., Batra, R., Cockfield, J. D., Wyncoll, D., et al. (2007). An outbreak in an intensive care unit of a strain of methicillin-resistant *Staphylococcus aureus* sequence type 239 associated with an increased rate of vascular access device-related bacteremia. *Clin. Infect. Dis.* 44, 493–501. doi: 10.1086/511034

El-baz, R., Rizk, D. E., Barwa, R., and Hassan, R. (2017). Virulence characteristics and molecular relatedness of methicillin resistant *Staphylococcus aureus* harboring different staphylococcal cassette chromosome mec. *Microb. Pathog.* 113, 385–395. doi: 10.1016/j.micpath.2017.11.021

Fatholahzadeh, B., Emaneini, M., Aligholi, M., Gilbert, G., Taherikalani, M., Jonaidi, N., et al. (2009). Molecular characterization of methicillin-resistant *Staphylococcus aureus* clones from a teaching hospital in Tehran. *Jpn. J. Infect. Dis.* 62, 309–311. Available online at: http://www0.nih.go.jp/JJID/62/309.html

Ghaznavi-Rad, E., Nor Shamsudin, M., Sekawi, Z., Khoon, L. Y., Aziz, M. N., Hamat, R. A., et al. (2010). Predominance and emergence of clones of hospital-acquired methicillin-resistant *Staphylococcus aureus* in Malaysia. *J. Clin. Microbiol.* 48, 867–872. doi: 10.1128/JCM.01112-09

Gostev, V., Kruglov, A., Kalinogorskaya, O., Dmitrenko, O., Khokhlova, O., Yamamoto, T., et al. (2017). Molecular epidemiology and antibiotic resistance of methicillin-resistant *Staphylococcus aureus* circulating in the Russian Federation. *Infect. Genet. Evol.* 53, 189–194. doi: 10.1016/j.meegid.2017.06.006

Harkins, C. P., Pichon, B., Doumith, M., Parkhill, J., Westh, H., Tomasz, A., et al. (2017). Methicillin-resistant *Staphylococcus aureus* emerged long before the introduction of methicillin into clinical practice. *Genome Biol.* 18:130. doi: 10.1186/s13059-017-1252-9

Harris, S. R., Feil, E. J., Holden, M. T., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395

Havaei, S. A., Azimian, A., Fazeli, H., Naderi, M., Ghazvini, K., Samiee, S. M., et al. (2013). Isolation of Asian endemic and livestock associated clones of methicillin resistant *Staphylococcus aureus* from ocular samples in Northeastern Iran. *Iran. J. Microbiol.* 5, 227–232.

Holden, M. T. G., Lindsay, J. A., Corton, C., Quail, M. A., Cockfield, J. D., Pathak, S., et al. (2009). Genome sequence of a recently emerged highly-transmissible, multi-antibiotic and antiseptic resistant, variant of methicillin-resistant *Staphylococcus aureus* (MRSA) sequence-type 239 (TW). *J. Bacteriol.* 192, 888–892. doi: 10.1128/JB.01255-09

Howden, B. P., Seemann, T., Harrison, P. F., McEvoy, C. R., Stanton, J. A., Rand, C. J., et al. (2010). Complete genome sequence of *Staphylococcus aureus* strain JKD6008, a ST239 clone of methicillin-resistant *Staphylococcus aureus* with intermediate level vancomycin resistance. *J. Bacteriol.* 192, 5848–5849. doi: 10.1128/JB.00951-10

Hsu, L.-Y., Harris, S. R., Chlebowicz, M. A., Lindsay, J. A., Koh, T.-H., Krishnan, P., et al. (2015). Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol.* 16:81. doi: 10.1186/s13059-015-0643-z

Hsu, L.-Y., Loomba-Chlebicka, N., Koh, Y.-L., Tan, T.-Y., Krishnan, P., Lin, R. T.-P., et al. (2007). Evolving EMRSA-15 epidemic in Singapore hospitals. *J. Med. Microbiol.* 56, 376–379. doi: 10.1099/jmm.0.46950-0

Humphreys, H., Carroll, J. D., Keane, C. T., Cafferkey, M. T., Pomeroy, H. M., and Coleman, D. C. (1990). Importation of methicillin-resistant *Staphylococcus aureus* from Baghdad to Dublin and subsequent nosocomial spread. *J. Hosp. Infect.* 15, 127–135. doi: 10.1016/0195-6701(90)90121-4

Ip, M., Yung, R. W. H., Ng, T. K., Luk, W. K., Tse, C., Hung, P., et al. (2005). Contemporary methicillin-resistant *Staphylococcus aureus* clones in Hong Kong. *J. Clin. Microbiol.* 43, 5069–5073. doi: 10.1128/JCM.43.10.5069-5073.2005

Ito, T., Katayama, Y., Asada, K., Mori, N., Tsutsumimoto, K., Tiensasitorn, C., et al. (2001). Structural comparison of three types of staphylococcal cassette chromosome mec integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 45, 1323–1336. doi: 10.1128/AAC.45.5.1323-1336.2001

Iwao, Y., Khokhlova, O. E., Takano, T., Hung, W. C., Isobe, H., Peryanova, O. V., et al. (2012). Fatal pneumonia in HIV-infected patients from a novel ST239

methicillin-resistant *Staphylococcus aureus* carrying the toxic shock syndrome toxin-1 gene in Krasnoyarsk, Siberian Russia. *Jpn. J. Infect. Dis.* 65, 184–186. Available online at: http://www.nih.go.jp/niid/images/JJID/65-2/184.pdf

IWG-SCC (2009). Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrob. Agents Chemother.* 53, 4961–4967. doi: 10.1128/AAC.00579-09

Jamil, B., Gawlik, D., Syed, M. A., Shah, A. A., Abbasi, S. A., Müller, E., et al. (2017). Hospital-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) from Pakistan: molecular characterisation by microarray technology. *Eur. J. Clin. Microbiol. Infect. Dis.* 37, 691–700. doi: 10.1007/s10096-017-3161-y

Jansen van Rensburg, M. J., Eliya Madikane, V., Whitelaw, A., Chachage, M., Haffejee, S., and Gay Elisha, B. (2011). The dominant methicillin-resistant *Staphylococcus aureus* clone from hospitals in Cape Town has an unusual genotype: ST612. *Clin. Microbiol. Infect.* 17, 785–792. doi: 10.1111/j.1469-0691.2010.03373.x

Khokhlova, O. E., Hung, W.-C., Wan, T.-W., Iwao, Y., Takano, T., Higuchi, W., et al. (2015). Healthcare- and community-associated methicillin-resistant *Staphylococcus aureus* (MRSA) and fatal pneumonia with pediatric deaths in Krasnoyarsk, Siberian Russia: unique MRSA's multiple virulence factors, genome, and stepwise evolution. *PLoS ONE* 10:e0128017. doi: 10.1371/journal.pone.0128017

Krziwanek, K., Luger, C., Sammer, B., Stumvoll, S., Stammler, M., Sagel, U., et al. (2008). MRSA in Austria-an overview. *Clin. Microbiol. Infect.* 14, 250–259. doi: 10.1111/j.1469-0691.2007.01896.x

Kwon, N. H., Park, K. T., Jung, W. K., Youn, H. Y., Lee, Y., Kim, S. H., et al. (2006). Characteristics of methicillin resistant *Staphylococcus aureus* isolated from chicken meat and hospitalized dogs in Korea and their epidemiological relatedness. *Vet. Microbiol.* 117, 304–312. doi: 10.1016/j.vetmic.2006.05.006

Li, M., Du, X., Villaruz, A. E., Diep, B. A., Wang, D., Song, Y., et al. (2012). MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat. Med.* 18, 816–819. doi: 10.1038/nm.2692

Luedicke, C., Slickers, P., Ehricht, R., and Monecke, S. (2010). Molecular fingerprinting of *Staphylococcus aureus* from bone and joint infections. *Eur. J. Clin. Microbiol. Infect. Dis.* 29, 457–463. doi: 10.1007/s10096-010-0884-4

Mayor, L., Ortellado, J., Menacho, C., Lird, G., Courtier, C., Gardon, C., et al. (2007). Molecular characterization of methicillin-resistant *Staphylococcus aureus* isolates collected in Asuncion, Paraguay. *J. Clin. Microbiol.* 45, 2298–2300. doi: 10.1128/JCM.00040-07

Melter, O., Aires de Sousa, M., Urbášková, P., Jakubu, V., Žemličková, H., and de Lencastre, H. (2003). Update on the major clonal types of methicillin-resistant *Staphylococcus aureus* in the Czech Republic. *J. Clin. Microbiol.* 41, 4998–5005. doi: 10.1128/JCM.41.11.4998-5005.2003

Monecke, S., Coombs, G., Shore, A. C., Coleman, D. C., Akpaka, P., Borg, M., et al. (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

Monecke, S., Jatzwauk, L., Muller, E., Nitschke, H., Pfohl, K., Slickers, P., et al. (2016). Diversity of SCCmec elements in *Staphylococcus aureus* as observed in South-Eastern Germany. *PLoS ONE* 11:e0162654. doi: 10.1371/journal.pone.0162654

Monecke, S., Jatzwauk, L., Weber, S., Slickers, P., and Ehricht, R. (2008a). DNA microarray-based genotyping of methicillin-resistant *Staphylococcus aureus* strains from Eastern Saxony. *Clin. Microbiol. Infect.* 14, 534–545. doi: 10.1111/j.1469-0691.2008.01986.x

Monecke, S., Muller, E., Buechler, J., Rejman, J., Stieber, B., Akpaka, P. E., et al. (2013a). Rapid detection of panton-valentine leukocidin in *Staphylococcus aureus* cultures by use of a lateral flow assay based on monoclonal antibodies. *J. Clin. Microbiol.* 51, 487–495. doi: 10.1128/JCM.02285-12

Monecke, S., Müller, E., Dorneanu, O. S., Vremeră, T., and Ehricht, R. (2014a). Molecular typing of MRSA and of clinical *Staphylococcus aureus* isolates from Iaşi, Romania. *PLoS ONE* 9:e97833. doi: 10.1371/journal.pone.0097833

Monecke, S., Muller, E., Schwarz, S., Hotzel, H., and Ehricht, R. (2012a). Rapid microarray based identification of different mecA alleles in Staphylococci. *Antimicrob. Agents Chemother.* 56, 5547–5554. doi: 10.1128/AAC.00574-12

Monecke, S., Nitschke, H., Slickers, P., Ehricht, R., Swanston, W., Manjunath, M., et al. (2012b). Molecular epidemiology and characterisation of MRSA isolates from Trinidad and Tobago. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 1497–1500. doi: 10.1007/s10096-011-1469-6

Monecke, S., Ruppelt, A., Wendlandt, S., Schwarz, S., Slickers, P., Ehricht, R., et al. (2013b). Genotyping of *Staphylococcus aureus* isolates from diseased poultry. *Vet. Microbiol.* 162, 806–812. doi: 10.1016/j.vetmic.2012.10.018

Monecke, S., Skakni, L., Hasan, R., Ruppelt, A., Ghazal, S. S., Hakawi, A., et al. (2012c). Characterisation of MRSA strains isolated from patients in a hospital in Riyadh, Kingdom of Saudi Arabia. *BMC Microbiol.* 12:146. doi: 10.1186/1471-2180-12-146

Monecke, S., Slickers, P., and Ehricht, R. (2008b). Assignment of *Staphylococcus aureus* isolates to clonal complexes based on microarray analysis and pattern recognition. *FEMS Immunol. Med. Microbiol.* 53, 237–251. doi: 10.1111/j.1574-695X.2008.00426.x

Monecke, S., Stieber, B., Roberts, R., Akpaka, P. E., Slickers, P., and Ehricht, R. (2014b). Population structure of *Staphylococcus aureus* from Trinidad & Tobago. *PLoS ONE* 9:e89120. doi: 10.1371/journal.pone.0089120

Neetu, T. J., and Murugan, S. (2016). Genotyping of methicillin resistant *Staphylococcus aureus* from Tertiary Care Hospitals in Coimbatore, South India. *J. Glob. Infect. Dis.* 8, 68–74. doi: 10.4103/0974-777X.182119

Nemati, M., Hermans, K., Lipinska, U., Denis, O., Deplano, A., Struelens, M., et al. (2008). Antimicrobial resistance of old and recent *Staphylococcus aureus* isolates from poultry: first detection of livestock-associated methicillin-resistant strain ST398. *Antimicrob. Agents Chemother.* 52, 3817–3819. doi: 10.1128/AAC.00613-08

Nimmo, G. R., Steen, J. A., Monecke, S., Ehricht, R., Slickers, P., Thomas, J. C., et al. (2015). ST2249-MRSA-III: a second major recombinant methicillin-resistant *Staphylococcus aureus* clone causing healthcare infection in the 1970s. *Clin. Microbiol. Infect.* 21, 444–450. doi: 10.1016/j.cmi.2014.12.018

Oliveira, D. C., Wu, S. W., and de Lencastre, H. (2000). Genetic organization of the downstream region of the mecA element in methicillin-resistant *Staphylococcus aureus* isolates carrying different polymorphisms of this region. *Antimicrob. Agents Chemother.* 44, 1906–1910. doi: 10.1128/AAC.44.7.1906-1910.2000

Orth, D., Grif, K., Erdenechimeg, L., Battogtokh, C., Hosbayar, T., Strommenger, B., et al. (2006). Characterization of methicillin-resistant *Staphylococcus aureus* from Ulaanbaatar, Mongolia. *Eur. J. Clin. Microbiol. Infect. Dis.* 25, 104–107. doi: 10.1007/s10096-006-0102-6

Peck, K. R., Baek, J. Y., Song, J. H., and Ko, K. S. (2009). Comparison of genotypes and enterotoxin genes between *Staphylococcus aureus* isolates from blood and nasal colonizers in a Korean hospital. *J. Korean Med. Sci.* 24, 585–591. doi: 10.3346/jkms.2009.24.4.585

Robinson, D., and Enright, M. (2004a). Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 10, 92–97. doi: 10.1111/j.1469-0691.2004.00768.x

Robinson, D. A., and Enright, M. C. (2004b). Evolution of *Staphylococcus aureus* by large chromosomal replacements. *J. Bacteriol.* 186, 1060–1064. doi: 10.1128/JB.186.4.1060-1064.2004

Rossney, A. S. (2003). *A New Strain of MRSA Imported from Singapore.* EPI-Insight, Disease Surveillance Report of NDSC, Ireland.

Rossney, A. S., Coleman, D. C., and Keane, C. T. (1994). Evaluation of an antibiogram-resistogram typing scheme for methicillin-resistant *Staphylococcus aureus*. *J. Med. Microbiol.* 41, 441–447. doi: 10.1099/00222615-41-6-441

Roy, A. F., Newman, S. S., Cox, H. U., and Hoskins, J. D. (1984). Effect of beta-lactamase of *Staphylococcus intermedius* on disk agar diffusion susceptibility tests. *Cornell Vet.* 74, 354–360.

Schaefler, S., Jones, D., Perry, W., Ruvinskaya, L., Baradet, T., Mayr, E., et al. (1981). Emergence of gentamicin- and methicillin-resistant *Staphylococcus aureus* strains in New York City hospitals. *J. Clin. Microbiol.* 13, 754–759.

Scicluna, E., Shore, A., Thuermer, A., Ehricht, R., Slickers, P., Borg, M., et al. (2010). Characterisation of MRSA from Malta and the description of a Maltese epidemic MRSA strain. *Eur. J. Clin. Microbiol. Infect. Dis.* 29, 163–170. doi: 10.1007/s10096-009-0834-1

Senok, A., Ehricht, R., Monecke, S., Al-Saedan, R., and Somily, A. (2016). Molecular characterization of methicillin-resistant *Staphylococcus aureus* in nosocomial infections in a tertiary-care facility: emergence of new clonal complexes in Saudi Arabia. *New Microbes New Infect* 14, 13–18. doi: 10.1016/j.nmni.2016.07.009

Shabir, S., Hardy, K. J., Abbasi, W. S., McMurray, C. L., Malik, S. A., Wattal, C., et al. (2010). Epidemiological typing of meticillin-resistant *Staphylococcus*

aureus isolates from Pakistan and India. *J. Med. Microbiol.* 59, 330–337. doi: 10.1099/jmm.0.014910-0

Shang, W., Hu, Q., Yuan, W., Cheng, H., Yang, J., Hu, Z., et al. (2016). Comparative fitness and determinants for the characteristic drug resistance of ST239-MRSA-III-t030 and ST239-MRSA-III-t037 strains isolated in China. *Microb. Drug Resist.* 22, 185–192. doi: 10.1089/mdr.2015.0226

Shittu, A., Nubel, U., Udo, E., Lin, J., and Gaogakwe, S. (2009). Characterization of meticillin-resistant *Staphylococcus aureus* isolates from hospitals in KwaZulu-Natal province, Republic of South Africa. *J. Med. Microbiol.* 58, 1219–1226. doi: 10.1099/jmm.0.011452-0

Shore, A., Rossney, A. S., Keane, C. T., Enright, M. C., and Coleman, D. C. (2005). Seven novel variants of the staphylococcal chromosomal cassette *mec* in methicillin-resistant *Staphylococcus aureus* isolates from Ireland. *Antimicrob. Agents Chemother.* 49, 2070–2083. doi: 10.1128/AAC.49.5.2070-2083.2005

Shore, A. C., and Coleman, D. C. (2013). Staphylococcal cassette chromosome *mec*: recent advances and new insights. *Int. J. Med. Microbiol.* 303, 350–359. doi: 10.1016/j.ijmm.2013.02.002

Smyth, D. S., McDougal, L. K., Gran, F. W., Manoharan, A., Enright, M. C., Song, J. H., et al. (2010). Population structure of a hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. *PLoS ONE* 5:e8582. doi: 10.1371/journal.pone.0008582

Takano, T., Saito, K., Teng, L. J., and Yamamoto, T. (2007). Spread of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) in hospitals in Taipei, Taiwan in 2005, and comparison of its drug resistance with previous hospital-acquired MRSA. *Microbiol. Immunol.* 51, 627–632. doi: 10.1111/j.1348-0421.2007.tb03949.x

Tekeli, A., Ocal, D. N., Ozmen, B. B., Karahan, Z. C., and Dolapci, I. (2016). Molecular Characterization of Methicillin-Resistant *Staphylococcus aureus* Bloodstream Isolates in a Turkish University Hospital Between 2002 and 2012. *Microb. Drug Resist.* 22, 564–569. doi: 10.1089/mdr.2015.0116

Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H., et al. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J. Clin. Microbiol.* 33, 2233–2239.

Udo, E. E., and Al-Sweih, N. (2017). Dominance of community-associated methicillin-resistant *Staphylococcus aureus* clones in a maternity hospital. *PLoS ONE* 12:e0179563. doi: 10.1371/journal.pone.0179563

van Belkum, A., Melles, D. C., Nouwen, J., van Leeuwen, W. B., van Wamel, W., Vos, M. C., et al. (2009). Co-evolutionary aspects of human colonisation and infection by *Staphylococcus aureus*. *Infect. Genet. Evol.* 9, 32–47. doi: 10.1016/j.meegid.2008.09.012

Vandenesch, F., Naimi, T., Enright, M., Lina, G., Nimmo, G., Heffernan, H., et al. (2003). Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. *Emerging Infect. Dis.* 9, 978–984. doi: 10.3201/eid0908.030089

Vivoni, A. M., Diep, B. A., de Gouveia Magalhaes, A. C., Santos, K. R., Riley, L. W., Sensabaugh, G. F., et al. (2006). Clonal composition of *Staphylococcus aureus* isolates at a Brazilian university hospital: identification

of international circulating lineages. *J. Clin. Microbiol.* 44, 1686–1691. doi: 10.1128/JCM.44.5.1686-1691.2006

Wang, Z., Zhou, H., Wang, H., Chen, H., Leung, K. K., Tsui, S., et al. (2014). Comparative genomics of methicillin-resistant *Staphylococcus aureus* ST239: distinct geographical variants in Beijing and Hong Kong. *BMC Genomics* 15:529. doi: 10.1186/1471-2164-15-529

Weber, S., Ehricht, R., Slickers, P., Abdel-Wareth, L., Donnelly, G., Pitout, M., et al. (2010). "Genetic fingerprinting of MRSA from Abu Dhabi," in *ECCMID 2010* (Vienna).

Witte, W., Kresken, M., Braulke, C., and Cuny, C. (1997). Increasing incidence and widespread dissemination of methicillin-resistant *Staphylococcus aureus* (MRSA) in hospitals in central Europe, with special reference to German hospitals. *Clin. Microbiol. Infect.* 3, 414–422. doi: 10.1111/j.1469-0691.1997.tb00277.x

Yamamoto, T., Takano, T., Higuchi, W., Iwao, Y., Singur, O., Reva, I., et al. (2012). Comparative genomics and drug resistance of a geographic variant of ST239 methicillin-resistant *Staphylococcus aureus* emerged in Russia. *PLoS ONE* 7:e29187. doi: 10.1371/annotation/05a851e1-8627-4ea4-b12a-cfdba09f1e61

Yeap, A. D., Woods, K., Dance, D. A. B., Pichon, B., Rattanavong, S., Davong, V., et al. (2017). Molecular epidemiology of *Staphylococcus aureus* skin and soft tissue infections in the Lao People's Democratic Republic. *Am. J. Trop. Med. Hyg.* 97, 423–428. doi: 10.4269/ajtmh.16-0746

Zafar, A., Stone, M., Ibrahim, S., Parveen, Z., Hasan, Z., Khan, E., et al. (2011). Prevalent genotypes of meticillin-resistant *Staphylococcus aureus*: report from Pakistan. *J. Med. Microbiol.* 60, 56–62. doi: 10.1099/jmm.0.022707-0

Zurita, J., Barba, P., Ortega-Paredes, D., Mora, M., and Rivadeneira, S. (2016). Local circulating clones of *Staphylococcus aureus* in Ecuador. *Braz. J. Infect. Dis.* 20, 525–533. doi: 10.1016/j.bjid.2016.08.006

Genomic Comparison of Highly Virulent, Moderately Virulent, and Avirulent Strains From a Genetically Closely-Related MRSA ST239 Sub-lineage Provides Insights Into Pathogenesis

Jo-Ann M. McClure[1], Sahreena Lakhundi[1,2], Ayesha Kashif[1], John M. Conly[1,2,3,4,5] and Kunyan Zhang[1,2,3,4,5]*

[1] *Centre for Antimicrobial Resistance, Alberta Health Services/Calgary Laboratory Services/University of Calgary, Calgary, AB, Canada,* [2] *Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, AB, Canada,* [3] *Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Calgary, AB, Canada,* [4] *Department of Medicine, University of Calgary, Calgary, AB, Canada,* [5] *The Calvin, Phoebe and Joan Snyder Institute for Chronic Diseases, University of Calgary, Calgary, AB, Canada*

The genomic comparison of virulent (TW20), moderately virulent (CMRSA6/CMRSA3), and avirulent (M92) strains from a genetically closely-related MRSA ST239 sub-lineage revealed striking similarities in their genomes and antibiotic resistance profiles, despite differences in virulence and pathogenicity. The main differences were in the *spa* gene (coding for staphylococcal protein A), *lpl* genes (coding for lipoprotein-like membrane proteins), *cta* genes (genes involved in heme synthesis), and the *dfrG* gene (coding for a trimethoprim-resistant dihydrofolate reductase), as well as variations in the presence or content of some prophages and plasmids, which could explain the virulence differences of these strains. TW20 was positive for all genetic traits tested, compared to CMRSA6, CMRSA3, and M92. The major components differing among these strains included *spa* and *lpl* with TW20 carrying both whereas CMRSA6/CMRSA3 carry *spa* identical to TW20 but have a disrupted *lpl*. M92 is devoid of both these traits. Considering the role played by these components in innate immunity and virulence, it is predicted that since TW20 has both the components intact and functional, these traits contribute to its pathogenesis. However, CMRSA6/CMRSA3 are missing one of these components, hence their intermediately virulent nature. On the contrary, M92 is completely devoid of both the *spa* and *lpl genes* and is avirulent. Mobile genetic elements play a potential role in virulence. TW20 carries three prophages (φSa6, φSa3, and φSPβ-like), a pathogenicity island and two plasmids. CMRSA6, CMRSA3, and M92 contain variations in one or more of these components. The virulence associated genes in these components include staphylokinase, entertoxins, antibiotic/antiseptic/heavy metal resistance and bacterial persistence. Additionally, there are many hypothetical proteins (present with variations among strains) with unknown function in these mobile elements which could

be making an important contribution in the virulence of these strains. The above mentioned repertoire of virulence components in TW20 likely contributes to its increased virulence, while the absence and/or modification of one or more of these components in CMRSA6/CMRSA3 and M92 likely affects the virulence of the strains.

INTRODUCTION

Methicillin-resistant *Staphylococcus aureus* (MRSA) continues to be a major cause of hospital infection, as well as an emerging cause of community associated infections (David and Daum, 2010; Kock et al., 2010). Enright et al. (2002) revealed that the majority of global MRSA clones belonged to one of the five major clonal complexes (CCs) including CC5, CC8 (including CC8-ST239 sub-group), CC22, CC30 and CC45, however, strains belonging to many other CCs are emerging as significant sources of infection. Within these CCs, ST239 carrying staphylococcal cassette chromosome *mec* (SCC*mec*) III, is a healthcare-associated MRSA lineage present worldwide (Aires de Sousa and de Lencastre, 2004; Harris et al., 2010; Gray et al., 2011; Hsu et al., 2015). ST239-MRSA-III is prevalent in Asia, South America and Eastern Europe, and includes strains like the Brazilian, Hungarian, Portuguese, AUS-EMRSA-2 and 3, Viennese and EMRSA-1, -4,-7, 9, and 11 clones (Aires de Sousa and de Lencastre, 2004; Conceicao et al., 2007; Monecke et al., 2011). Phylogenetic analysis has revealed the intercontinental dissemination and hospital transmission of CC8-ST239 isolates throughout North America, Europe, South America, and Asia (Harris et al., 2010; Wang et al., 2012). In the 1990s, ST239 dispersed from South America to Europe and from Thailand to China (Gray et al., 2011).

Holden et al. (2010) isolated a highly transmissible outbreak MRSA strain, TW20, belonging to ST239-MRSA-III from an intensive care unit (ICU) in a London hospital. TW20 was found to be a highly invasive MRSA strain, with its acquisition four times more likely to result in bacteremia as compared to the other epidemic strains like EMRSA-15 or -16. In addition, it was more frequently isolated from vascular device cultures, had an extended antibiotic resistance pattern and an elevated minimum bactericidal concentration for chlorhexidine compared to other MRSA strains. TW20 was consequently classified as a highly virulent MRSA strain. Interestingly, the association of this strain with carriage/colonization sites like the nares, axilla and perineum was less frequent, suggesting a difference in its colonization capability as compared to other MRSA strains (Holden et al., 2010). The authors investigated the genetic basis for this increased transmissibility, resistance and virulence by analyzing its whole genome and comparing it with other MRSA lineages. They identified two large mobile genetic regions, a wide range of genes responsible for antibiotic, antiseptic and heavy metal resistance, as well as mutations in some housekeeping genes, which could all be responsible for its increased virulence, invasiveness

and survival in the hospital environment (Holden et al., 2010).

In Canada, we have had several epidemic hospital-associated MRSA (HA-MRSA) strains (**Figure 1A**) identified. These include CMRSA6 and CMRSA3 (ST239-t037-MRSA-III and ST241-t037-MRSA-III respectively which are similar to USA epidemic pulsotype strain USA700) which before the emergence of community-associated MRSA (CA-MRSA), along with CMRSA2 (ST5-t002-MRSA-II, similar to USA100/800), were the dominant causes of healthcare-associated MRSA infections in Canada. CMRSA2 was the predominantly isolated HA-MRSA strain during 2000–2006, while CMRSA6 (replacing CMRSA3) was the most common strain isolated in the hospital in later years (Christianson et al., 2007). Compared to CMRSA2, infections caused by CMRSA6 and CMRSA3 were less frequent and less severe and/or invasive (Christianson et al., 2007). In a *Caenorhabditis elegans* MRSA virulence infection model, CMRSA6 and CMRSA3 killed only a small percentage of worms compared to CMRSA2, thereby classifying them as non- or low-nematocidal MRSA strains, which correlated well with clinically invasive anatomic site data from another study (Wu et al., 2010). CMRSA6 and CMRSA3 are therefore considered as moderately virulent strains.

In the late 1980s, M92, a colonizing MRSA strain and close relative of CMRSA6 and CMRSA3, was isolated from a hospital site in Calgary, AB, Canada. Over the course of many years, M92 was frequently found to be associated with nasal colonization in hospital staff and patients but was never found associated with infections (Wu et al., 2010; McClure and Zhang, 2017). In the *C. elegans* model, this benign strain did not show any nematocidal activity and was, therefore, classed as an avirulent strain and used as a control in many infection models (Wu et al., 2010, 2012a,b).

A whole genome comparison of TW20, CMRSA6, CMRSA3, and M92 revealed striking similarities. Sequence alignments of the genomes indicated that they are very closely related with only minor differences, which contrasts significantly to the virulence of these strains in clinical scenarios, with TW20 being highly virulent, CMRSA6 and CMRSA3, moderately virulent and M92, avirulent. While the differences are few, they likely represent genetic components which impact pathogenesis and could explain the virulence differences observed among the strains, both *in vivo* and in the clinical setting. Here we present whole genome sequence (WGS) comparisons among TW20, CMRSA6, CMRSA3 and M92, highlighting the factors that could play a role in the virulence and pathogenicity differences noted in these strikingly similar strains.
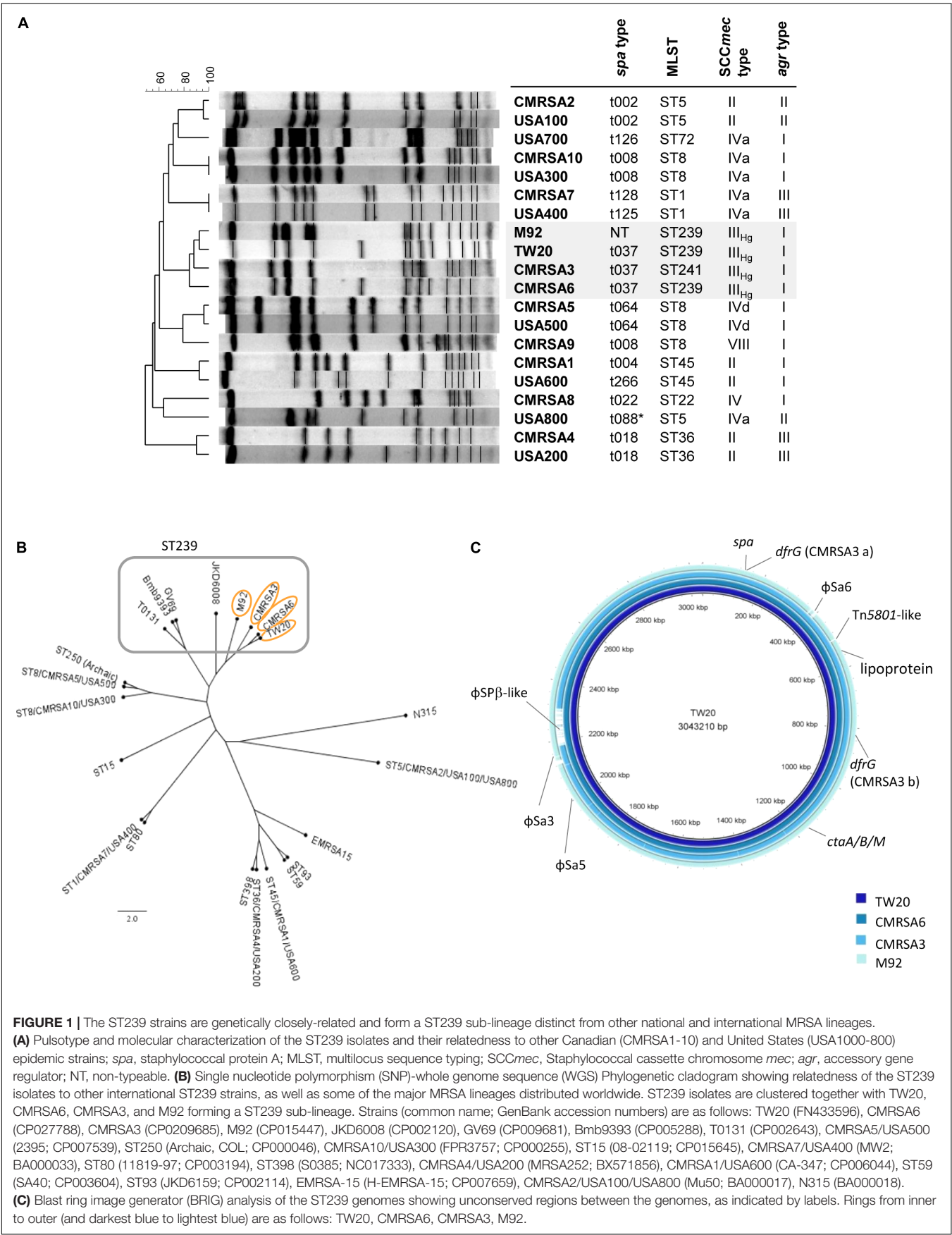
FIGURE 1 | The ST239 strains are genetically closely-related and form a ST239 sub-lineage distinct from other national and international MRSA lineages. **(A)** Pulsotype and molecular characterization of the ST239 isolates and their relatedness to other Canadian (CMRSA1-10) and United States (USA1000-800) epidemic strains; *spa*, staphylococcal protein A; MLST, multilocus sequence typing; SCC*mec*, Staphylococcal cassette chromosome *mec*; *agr*, accessory gene regulator; NT, non-typeable. **(B)** Single nucleotide polymorphism (SNP)-whole genome sequence (WGS) Phylogenetic cladogram showing relatedness of the ST239 isolates to other international ST239 strains, as well as some of the major MRSA lineages distributed worldwide. ST239 isolates are clustered together with TW20, CMRSA6, CMRSA3, and M92 forming a ST239 sub-lineage. Strains (common name; GenBank accession numbers) are as follows: TW20 (FN433596), CMRSA6 (CP027788), CMRSA3 (CP0209685), M92 (CP015447), JKD6008 (CP002120), GV69 (CP009681), Bmb9393 (CP005288), T0131 (CP002643), CMRSA5/USA500 (2395; CP007539), ST250 (Archaic, COL; CP000046), CMRSA10/USA300 (FPR3757; CP000255), ST15 (08-02119; CP015645), CMRSA7/USA400 (MW2; BA000033), ST80 (11819-97; CP003194), ST398 (S0385; NC017333), CMRSA4/USA200 (MRSA252; BX571856), CMRSA1/USA600 (CA-347; CP006044), ST59 (SA40; CP003604), ST93 (JKD6159; CP002114), EMRSA-15 (H-EMRSA-15; CP007659), CMRSA2/USA100/USA800 (Mu50; BA000017), N315 (BA000018). **(C)** Blast ring image generator (BRIG) analysis of the ST239 genomes showing unconserved regions between the genomes, as indicated by labels. Rings from inner to outer (and darkest blue to lightest blue) are as follows: TW20, CMRSA6, CMRSA3, M92.

MATERIALS AND METHODS

Bacterial Strains

The Canadian epidemic MRSA reference strains CMRSA1 to 10 (including CMRSA6 and CMRSA3) were provided by the National Microbiology Laboratory, Health Canada, Winnipeg, MB, Canada. The United States epidemic MRSA reference strains USA100 to USA800 (NRS382, NRS383, NRS384, NRS123, NRS385, NRS22, NRS386, and NRS387, respectively) were obtained through the Network on Antimicrobial Resistance in *Staphylococcus aureus* Program (NARSA) supported under NIAID/NIH contract no. N01-AI-95359. Strain M92 was kindly provided by Dr. T. Louie at University of Calgary, Canada, and strain TW20 by Dr. Julian Parkhill at the The Wellcome Trust Sanger Institute, United Kingdom.

Strain Molecular and Phenotypic Characterization

Staphylococcal isolates were fingerprinted by pulsed field gel electrophoresis (PFGE) after digestion with *Sma*I following a standardized protocol (Mulvey et al., 2001). PFGE-generated DNA fingerprints were digitized and analyzed with BioNumerics Ver. 6.6 (Applied Maths, Sint-Martens-Latem, Belgium) by using a position tolerance of 1.0 and an optimization of 1.0. Isolates were further characterized with Staphylococcal protein A (*spa*) typing (Harmsen et al., 2003), multilocus sequence typing (MLST) (Enright et al., 2000), SCC*mec* typing (McClure et al., 2010; Zhang et al., 2012), and accessory gene regulator (*agr*) typing (Peacock et al., 2002). Screening for antibiotic-resistant phenotypes was performed by use of VITEK 1 (bioMerieux) and the Clinical and Laboratory Standards Institute oxacillin agar screen, while confirmation of methicillin resistance was achieved using an in-house polymerase chain reaction (PCR) assay for the *mecA* gene (McClure et al., 2010; Zhang et al., 2012). Antibiotic susceptibility data for strain TW20 was obtained from the Holden study (Holden et al., 2010).

DNA Sequencing and Whole Genome Sequence Analysis

Genomic DNA for strains CMRSA6, CMRSA3, and M92 was isolated by phenol:chloroform extraction and sequenced with Pacific Biosciences (PacBio) RSII sequencing technology (McGill University Génome Québec Innovation Centre), as well as with Illumina MiSeq technology (Core DNA sequencing services, University of Calgary). Hybrid sequence assembly was performed using both read sets and the genomes annotated with NCBI's Prokaryotic Genomes Annotation Pipeline and deposited under accession numbers CP027788, CP029685, and CP015447, respectively. The genomes of TW20 were available under accession numbers FN433596 and CP015447, and the *spa* gene sequence of *S. aureus* strain 8325-4 was available under accession number J01786. Single nucleotide polymorphism (SNP) WGS phylogenetic analysis was performed using CSI Phylogeny v1.4 with default settings, using strain N315 (BA000018) as the reference and rooting genome (Center for Genomic Epidemiology, Kongens Lyngby, Denmark). Phylogenetic trees

were visualized with FigTree v1.4.3 (Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom). Blast ring images were generated using BRIG v0.95 (Alikhan et al., 2011). Representative genomic structures were generated and analyzed with Vector NTI Advance v11.5.2 (Invitrogen), DNA multiple sequence alignment with TCoffee (Notredame et al., 2000; Wallace et al., 2006; Moretti et al., 2007; Di Tommaso et al., 2011), and protein translation using fr33.net (France). Prophage identification and annotation was conducted using PHASTER software (Zhou et al., 2011; Arndt et al., 2016), and the comparisons using Easyfig (Sullivan et al., 2011).

RESULTS AND DISCUSSION

TW20, CMRSA6, CMRSA3, and M92 Form a Genetically Closely-Related ST239 Sub-lineage

Pulsed field gel electrophoresis analysis of TW20 revealed that it clustered together with CMRSA6, CMRSA3, and M92 when compared against major United States/Canada epidemic strains (**Figure 1A**). Molecular characterization of TW20, CMRSA6, CMRSA3, and M92 showed that they all carry SCC*mec* type III$_{Hg}$ and *agr* type I (**Figure 1A**), as well as the same *spa* type (t037), with the exception of M92 which is non-typeable via *spa* typing (**Figure 1A**). The strains belong to ST239, a major dominant hospital associated MLST type, except for CMRSA3 which belongs to the closely related ST241. ST241 differs from ST239 by a single point mutation at 268 bp in the *yqiL* locus (coding for acetyl coenzyme A acetyltransferase) whereby adenine in ST239 is replaced by a guanine (A to G). CMRSA3 is consequently very closely related to TW20, CMRSA6, and M92 (**Figure 1A**).

The genome of TW20 was published and is available (Holden et al., 2010), therefore whole genome sequencing was done on CMRSA6, CMRSA3, and M92 in order to analyze and compare the genomes of all four strains of this ST239 sub-lineage. The genome of TW20 is 3.0 Mbp and is reported to be the largest sequenced *S. aureus* genome. The genome of CMRSA6, CMRSA3, and M92 are similar in size, being 3.0, 2.9, and 3.0 Mbp, respectively. SNP whole genome phylogenetic analysis of TW20, CMRSA6, CMRSA3, and M92 supported the close genetic relatedness of these strains. They clustered apart from the other dominant global MRSA lineages (such as ST59 EMRSA-15, and USA300) and clustered together with international ST239 isolates (like GV69, T0131, and JKD6008) but forming a distinct sub-lineage within the ST239 (**Figure 1B**).

Detailed analysis of the TW20 genome was performed by Holden et al. (2010), and they suggested that several components in the genome could play a role in its highly virulent nature. These components included two large regions of 635 and 127 kb, as well as genes coding for QacA (antiseptic resistance protein), CadA (cadmium-transporting ATPase), TetM (tetracycline resistance protein), and DfrG (trimethoprim-resistant dihydrofolate reductase). Genes carried on mobile genetic elements like prophages (such as SCIN, *sek* and *sea*) or a pathogenicity island (such as *entK*, *entQ*) were also suspected

as being responsible for the high virulence. A surface anchored protein with the LPxTG motif, as well as point mutations in housekeeping genes coding for DNA gyrase subunit A (Ser84Leu; shown to confer resistance to quinolones) and isoleucyl-tRNA synthetase (Val588Phe; shown to confer low level mupirocin resistance), were also mentioned.

Comparisons were made between the genomes of the highly virulent TW20, the moderately virulent CMRSA6 and CMRSA3, and the avirulent M92, with respect to all regions mentioned by Holden et al. (2010). Careful analysis revealed that the above mentioned elements were, for the most part, present and similar among the four strains and, therefore, less likely to account for the increased virulence of TW20. Our analysis of the four genomes did reveal regions that differed between the ST239 strains, and these regions could be responsible for the elevated virulence of TW20. Differences were noted in *spa* (coding for staphylococcal protein A, SpA, a wall anchored protein imparting *S. aureus* with the ability to avoid opsonins present in normal serum), the *lpl* gene (coding for lipoprotein-like membrane proteins with N-terminal lipid moiety anchoring it to the outer leaflet of the cytoplasmic membrane), the *cta* genes (genes involved in heme synthesis), which were not identified as the virulent factors in Holden's study (Holden et al., 2010), as well as, *dfrG*/conserved hypothetical protein genes (coding for trimethoprim-resistant form of dihydrofolate reductase), which was pointed out by Holden et al. (2010) (**Figure 1C**). In addition, there were variations in the presence or content of prophages, including φSa6/5 (φSa6 formerly classified as φSa1 in TW20), φSa3 and φSPβ-like (**Figure 1C**). Furthermore, CMRSA3 and TW20 carried plasmids which contained similar content and were found to be missing in CMRSA6 and M92. Each of these unique regions within the chromosome was studied in detail to understand its possible role in the virulence and pathogenicity in *S. aureus*.

spa Gene Truncation and SpA Functional Destitution in M92

SpA, encoded by *spa*, plays an important role in the virulence and pathogenicity of *S. aureus* through innate immune evasion. The prototype *spa* gene from *S. aureus* strain 8325-4 contains an open reading frame (ORF) consisting of 1576 bp, giving rise to a protein of $M_r = 58,703$ (Uhlen et al., 1984), and is considered to be a complete *spa* gene. As seen in **Figure 2**, SpA of strain 8325-4 is composed of multiple domains, including the signal peptide (S) at the N terminus, responsible for directing the protein to its destination, followed by five highly homologous repeated domains (E, D, A, B, and C), which correspond to the immunoglobulin (Ig) binding domains and are found to vary in different *S. aureus* strains. The Xr region (also called the cell wall binding domain) follows, consisting of variable numbers of repeats, approximately eight amino acids in length, which form the basis of *spa* typing. Finally there is the X_C region, which is a transmembrane domain consisting of a LPxTG motif, a hydrophobic region and a charged tail (Schneewind et al., 1993). It plays an important role in anchoring the protein to the surface of the cell (Schneewind et al., 1992, 1993).

The *spa* genes of TW20 (SATW20_01230), CMRSA6 and CMRSA3 are identical to each other at 1,283 bp in length. They contain complete E, B, C and Xc domains, identical to those of 8325-4 (**Figure 2**). However, in all three strains the D and A domains are partially missing (from 342 to 515 bp) (**Figure 2**). TW20, CMRSA6 and CMRSA3 also differ from 8325-4 in their Xr regions, both in terms of sequence composition and length (184 bp as compared to the 289 bp of strain 8325-4), which forms the basis for their differential *spa* classifications. In strain M92, by contrast, the majority of the *spa* gene is truncated, leaving a gene of 425 bp in length. Domains S and E are present, while domains A, B, and C are completely missing and only 25 bp of domain D is present, likely making it non-functional (**Figure 2**). While the E domain is present in M92 *spa* gene, its role as an Ig binding domain is controversial. Earlier studies have indicated that domain E has diverged more than the other four domains (A–D) and therefore probably has a different biological function instead of Ig binding (Hjelm et al., 1975; Sjodahl, 1977a,b; Wright et al., 1977; Hanson and Schumaker, 1984; Uhlen et al., 1984). Domain Xr is completely missing in M92 making it non-typeable via *spa* typing. The majority of the N-terminal part (168 bp out of 291 bp) of domain Xc is missing as well (**Figure 2**). This means that the LPxTG motif, hydrophobic domain and probably a part of charged tail are missing from M92 *spa* and the protein may not anchor to the surface of the cell. Thus, M92 *spa* may potentially be devoid of any function of SpA.

SpA can be regarded as an innate immune evasion molecule, conferring the ability to survive within the host and cause successful infection, and has been shown to be present in 98% of coagulase positive *S. aureus* strains (Forsgren, 1970). The immunoglobulin binding domain of SpA has the affinity to bind to both the Fc portion of IgG, as well as the Fab portion of the V_H3 region of IgM located on the surface of B cells. Because SpA normally resides on the organism's cell surface, the interaction of IgG with SpA results in coating of the pathogen's cell surface with IgG molecules. These IgG molecules are in the incorrect orientation for recognition by the Fc receptors of neutrophils, inhibiting opsonophagocytic killing of the organism (Foster, 2005; Rooijakkers et al., 2005). On the other hand, the ability of SpA to bind to the Fab portion of the V_H3 region of IgM on the surface of B cells causes the cells to proliferate and undergo apoptosis. This diminishes the repertoire of antibody-secreting B lymphocytes in the spleen and bone marrow (Goodyear and Silverman, 2004). As mentioned, the SpA of TW20, CMRSA6, and CMRSA3 contain complete immunoglobulin binding domains B, C, and potentially E. Although domains A and D are partial and likely non-functional, these strains still have 2 or 3 complete Fc and Fab binding regions, playing a role in protecting these strains from the host immune system. In contrast, the SpA of M92 is completely missing domains A, B and C, with only 25 bp of domain D present and is likely non-functional. While domain E is present, its controversial role as an Ig binding domain possibly leaves M92 *spa* devoid of any IgG binding domains. In addition, due to the lack of a LPxTG motif, SpA from M92 is probably not expressed on the cell surface. As a consequence, once inside the human body, a lack of functional SpA on M92's cell surface

FIGURE 2 | Structural and sequence comparison of the *spa* genes among the ST239 isolates shows deletions in TW20, CMRSA6, CMRSA3, and M92. Structural arrangement of the complete *spa* gene in strain 8325-4 (Accession number J01786) showing regions coding for the signal peptidase domain (S), IgG binding domains (E, D, A, B, C), the short sequence repeat region (Xr), and the anchoring domain (Xc). Deletions in the *spa* gene of TW20, CMRSA6, CMRSA3, and M92 are mapped and indicated with a dotted line. The Xr regions differ between the strains and is the basis for their assignment to different *spa* types. Nucleotide positions in 8325-4 are indicated at the top.

FIGURE 3 | The lipoprotein sequence of TW20 is compared to the tandem lipoproteins in CMRSA6, CMRSA3, and M92. Structure of the TW20 lipoprotein gene as compared to the tandem lipoprotein genes in CMRSA6, CMRSA3 and M92, which are identical and contain a 1,679 bp insertion. The location of insertion relative to the TW20 sequence is marked with a line and falls between base pairs 135 and 136. Numbers indicate nucleotide locations within the DNA sequence, while numbers in brackets for CMRSA6, CMRSA3, and M92 represent corresponding nucleotides in the TW20 gene.

might result in its opsonization and killing by human immune cells.

lpl Disruption in CMRSA6, CMRSA3 and M92, but Not in TW20

TW20, CMRSA6, CMRSA3 and M92, all belong to the clonal complex CC8, most of which carry a conserved genomic island, *v*Sa*α* (Babu et al., 2006). This genomic island is characterized by two clusters of tandem repeat sequences, including an exotoxin (*set*) and a number of homologous lipoproteins (Lpp) arranged in tandem and referred to as lipoprotein-like (*lpl*) (Babu et al., 2006; Baba et al., 2008; Tsuru and Kobayashi, 2008). The exact function of these lipoproteins (Lpl)is not known, however they have recently been shown to trigger host cell invasion, increase pathogenicity and may contribute to the epidemic nature of CC8 and CC5 strains (Nguyen et al., 2015).

The *lpl* of TW20 (SATW20_05130) is a 816 bp sequence which results in the transcription of a protein with 271 amino acids (**Figure 3**). The transcribed protein from TW20 resembles Lpl and proteins containing conserved motifs called DUF567 (domain of unknown function). The *lpl* of CMRSA6, CMRSA3 and M92, in contrast, have 1,679 bp of DNA inserted within *lpl* (**Figure 3**), resulting in a gene size of 2,491 bp. The insertion is at bp 136 in the TW20 gene (as indicated by a line in **Figure 3**) and is possibly due to homologous recombination. The first 136 bp of sequence in TW20, CMRSA6, CMRSA3 and M92 are all identical, as are the next 36 bps (sequence: GAACAAATCAAAAAGAGCTTTGCGAAAACATTAGAT) of both the normal TW20 *lpl* and the inserted DNA. This 36 bp sequence may represent a region where homologous recombination occurred, permitting incorporation of the extra DNA into the *lpl* genes of CMRSA6, CMRSA3, and M92. Regardless of the mechanism involved, incorporation of the extra

FIGURE 4 | *ctaA*, *ctaB*, and *ctaM* sequence comparison among the ST239 isolates showing an IS256 insertion in CMRSA6. The structural arrangement of the *ctaA/B/M* genes is identical in strains TW20, CMRSA3 and M92, while CMRSA6 shows a 1,332 bp insertion of the IS256 transposase between the *ctaA* and *ctaB* genes. Nucleotide positions of the region are noted above, with the location of the insertion marked by a line at position 1,253 bp. The locations of the repeat sequence "tttttctct" are indicated.

DNA results in disruption of the original *lpl* gene. Fortunately, the interruption continues in the same ORF as the original gene and transcription of the gene in CMRSA6, CMRSA3, and M92 results in a new Lpl. This new Lpl is 271 amino acids long, but is missing a large part of the C-terminus of the original *lpl*, which may have functioned as a cell wall anchor. Interestingly, following the new Lpl lies a second ORF of 488 bp, which could code for a protein of 161 amino acids. However, this frame does not contain a start codon and is therefore unlikely to be transcribed into a protein.

Of interest is the fact that the genomes of some virulent strains, like Newman, can contain both types of *lpl* sequences; the original one similar to TW20, as well as the one with the insertion similar to CMRSA6, CMRSA3, and M92. We did not, however, find the inserted (1,679 bp) sequence anywhere within the TW20 genome, nor was the original *lpl* without insertion found in the genomes of the other three strains. It is also interesting to note that Lpl and the proteins containing DUF567 have been found to be taxonomically restricted to staphylococci and have recently been shown to play a significant role in the pathogenicity and virulence of *S. aureus* USA300 (Nguyen et al., 2015; Shahmirzadi et al., 2016).

Nguyen et al. (2015) deleted the entire *lpl* gene cluster in *S. aureus* USA300, making the mutant strain less invasive, with decreased ability to stimulate pro-inflammatory cytokines compared to the original or complemented strain. Invasiveness helps a pathogen shield itself from the harmful effect of antimicrobials, as well as from the human immune system, thereby contributing to its virulence and pathogenesis. TW20 with its intact Lpl may have an increased ability to stimulate the production of pro-inflammatory cytokines and may have increased invasiveness, as compared to CMRSA6, CMRSA3 and M92 with their interrupted *lpl* genes.

Potential Disruption of *Cta* in CMRSA6

In gram positive bacteria, heme synthesis is an important pathway providing substrate for the production of terminal oxidases (Mogi et al., 1994). In a heme-iron deficient environment, *S. aureus* fulfills its iron requirement via a complex pathway involving several genes (*cta*) coding for enzymes required for the synthesis of heme A (Hammer et al., 2016). The *cta* genes include *ctaA* (911 bp, coding for CtaA of 302 amino acids), *ctaB* (912 bp, coding for CtaB of 303 amino acids), and *ctaM* (463 bp, coding for CtaM of 153 amnio acids) (**Figure 4**). The orientation of *ctaA* and *ctaB* is in opposite direction to each other, with 441 bp between them, while *ctaM* is oriented in the same direction as *ctaB* (**Figure 4**). Whole genome analysis of TW20, CMRSA3, and M92 revealed that the *cta* genes in these strains are identical. The *cta* genes of CMRSA6, in contrast, have an IS256 insertion of 1,332 bp within the 441 bp region between *ctaA* and *ctaB*, with characteristic repeat regions (TTTTCTCT) at 1,253 bp (**Figure 4**). It is noteworthy that this insertion did not disrupt any gene, however, we are not sure if it resulted in the disruption of promotor function for any *cta* genes, as no promotor information is available. This promotor or promotors, if present, may be controlling the transcription of a single, or multiple *cta* genes, meaning that disruption in promotor function could potentially lead to the loss of the expression of one or multiple *cta* genes. All these genes play an important role in heme synthesis; Heme B is converted to heme O via CtaB which is then converted to heme A via CtaA (Svensson et al., 1993; Svensson and Hederstedt, 1994; Clements et al., 1999). CtaM was recently shown to support the function of QoxABCD, a respiratory oxygen reductase (Hammer et al., 2016). Therefore, the loss of even a single gene expression might affect the function of other genes.

Several studies have demonstrated that mutations in the *cta* genes results in decreased ability of the organism to survive long term starvation (Clements et al., 1999), decreased pigment production, attenuation of hemolytic activity and decreased growth (Lan et al., 2010; Xu et al., 2016). It also resulted in decreased transcription of several virulence genes, thereby affecting virulence (Xu et al., 2016), as well as host specific organ colonization of the organism (Hammer et al., 2013, 2016).

FIGURE 5 | Insertion of the trimethoprim resistance gene, *dfrG*, in ST239 strains. **(A)** Partial Tn*5801*-like element structure demonstrating insertion of the *dfrG* gene, plus 2 genes coding for hypothetical proteins (colored black), in strains TW20 and CMRSA6. The insertion is absent in CMRSA3 and M92. Open reading frames (ORFs) surrounding the insertion site are indicated in gray. **(B)** CMRSA3 contains two copies of the *dfrG* and hypothetical protein genes, but located in two separate locations on the chromosome. ORFs surrounding the insertions are indicated by gray hatch. Protein products of each gene are listed.

Presence of DfrG, Conferring Trimethoprim Resistance, in TW20, CMRSA6, and CMRSA3

DfrG, a trimethoprim-resistant dihydrofolate reductase, confers resistance to the antibiotic trimethoprim, used for the treatment of *S. aureus* infections (Rouch et al., 1989). The genome of TW20 and CMRSA6 carry a 31.3 kb region of Tn*5801*-like

element (**Figure 5A**), which is similar to transposons ICEs (integrative and conjugative elements) found in the genome of other *S. aureus* strains like Mu50 (Kuroda et al., 2001) and Mu3 (Neoh et al., 2008). The Tn*5801*-like element is responsible for dissemination of the tetracycline resistance gene, *tetM*, which codes for a ribosomal protection protein conferring resistance to the action of tetracycline (de Vries et al., 2016). Within this Tn*5801*-like element in TW20

and CMRSA6, three additional genes have been observed, including *dfrG* (SATW20_04710), coding for trimethoprim-resistant dihydrofolate reductase, as well as two others coding for hypothetical proteins (**Figure 5A**). *dfrG* is 498 bp in length, coding for a protein of 165 amino acids long, while the other two genes are 1953 and 481 bp long, coding for a protein sequences of 650 and 159 amino acids in length (**Figure 5A**).

The genomes of CMRSA3 and M92 also carry the Tn*5801*-like element, however, the element in these strains is missing the three additional genes (*dfrG* and two hypothetical proteins) (**Figure 5A**). Interestingly, the genome of CMRSA3 still carries the genes for DfrG and both hypothetical proteins, but not associated with the Tn*5801* like element (**Figure 5B**). In fact, CMRSA3 carries two copies of the *dfrG* gene located at two different positions within its genome (**Figure 1C**). One copy is located adjacent to the 3′ end of the *spa* gene and is accompanied by both of the hypothetical proteins (CMRSA3a), while the other copy is found at approximately 810 kb and only accompanied by the larger (1953 bp) hypothetical protein (CMRSA3b) (**Figure 5B**). The copy of *dfrG*/hypothetical protein present at 810 kb is oriented in the same direction as the one in TW20 and CMRSA6, but the copy of *dfrG*/hypothetical proteins positioned near *spa* is oriented in the opposite direction to that in TW20 and CMRSA6 (**Figure 5B**).

Trimethoprim is an important antibiotic used for the treatment of staphylococcal infections, particularly skin and soft tissue infections (Rouch et al., 1989; Nathwani et al., 2008; Stevens et al., 2014). It acts by inhibiting an enzyme (dihydrofolate reductase) involved in the folate synthesis pathway (Miovic and Pizer, 1971). TW20, CMRSA6, and CMRSA3 all carry this trimethoprim-resistant dihydrofolate reductase enzyme and were phenotypically resistant to trimethoprim (**Table 1**), which is likely one of the factors contributing to survival and persistence of these strains in the high antibiotic selective pressures seen in hospitals. M92, devoid of this enzyme, is susceptible to the action of trimethoprim, which is confirmed in the antibiotic resistance profiles of the strains (**Table 1**). The function and significance of the hypothetical proteins present next to *dfrG* are unknown.

Prophages and Mobile Genetic Elements Present in the ST239 Isolates

The staphylococcal genome displays several large sequence blocks with high variability which can carry determinants for antibiotic resistance and/or virulence. These variable regions can be classified as prophages, pathogenicity islands or staphylococcal cassette chromosomes (Baba et al., 2008), and several of them are present in the genomes of all four ST239 isolates studied here. Using the online phage search tool, PHASTER, mobile genetic elements such as prophages, phage-like proteins and pathogenicity islands were located and annotated in each genome. Analysis of the TW20 genome revealed the presence of five prophages, phage-like proteins and/or pathogenicity islands, including ϕSa1, ϕSa3, ϕSPβ-like, SPβ-like proteins, SaPI1. ϕSa1 has been re-classified as ϕSa6 due to both the nature of its integrase, as well as the location of integration

(Kahankova et al., 2010; Hyman et al., 2012). The genome of CMRSA6 contains all five of those mobile genetic elements (ϕSa6, ϕSa3, ϕSPβ-like, SPβ-like proteins, SaPI1), while the genome of CMRSA3 contains three (ϕSa6, ϕSa3, SaPI1), and the genome of M92 contains five (ϕSa5, ϕSa3, ϕSPβ-like, SPβ-like proteins, SaPI1) (**Figure 6A**). With some exceptions, the mobile elements were generally quite similar from strain to strain.

ϕSa6 was found in the genomes of TW20, CMRSA6 and CMRSA3, while a very similar phage, ϕSa5, was found in the genome of M92. The integration site for ϕSa5 is located on the opposite side of the M92 genome (near 2000 kb) as the integration site for ϕSa6 in the other three genomes (near 400 kb) (**Figure 6A**). ϕSa5 also differs in that it is inserted in the opposite orientation as compared to ϕSa6 (**Figure 6B**). The same prophage attachment sequences were identified at the left and right extremities of ϕSa6 in TW20, CMRSA6 and CMRSA3 (attL/attR: AAAAAAGGGCAGA), however multiple attachment sequences were identified at each of the extremities of ϕSa5 in M92 (Supplementary Table 1). These include an external sequence pair (attR2/attL2: CTTTTTAAAATTA), and an internal sequence pair (attR3/attL3: TAATTTAGTTAT). This finding suggests that ϕSa5 of M92 may be a composite of two prophages, possibly created by the incomplete excision of one phage, followed by insertion of a second phage. A comparison of ϕSa6 from TW20, CMRSA6, and CMRSA3 revealed that they are highly similar in their protein content (**Figure 6B** and Supplementary Table 1). While ϕSa5 is similar to ϕSa6 in a proportion of the proteins, it differs in the integrase, some of the DNA metabolism proteins, and in the proteins involved with the portal, head, and tail. ϕSa5 also contains 12 phage related proteins between attR2 and attR3 which are not present in ϕSa6. Despite their similarities and differences, it is important to note that ϕSa6 and ϕSa5 are not carrying any known virulence factors which could be responsible for the differential virulence noted among these four strains.

A second prophage, ϕSa3, was found in the genomes of all four strains (TW20, CMRSA6, CMRSA3, and M92) (**Figure 6A**), present at the same location (near 2120 kb) in each strain, with the identical attachment sites (attL/R: AAGTTGCAACAC) identified in TW20, CMRSA6, and CMRSA3 (**Figures 6A,B** and Supplementary Table 2). M92, on the other hand, shares an identical attL sequence with the other three strains, however, it is located internal to alternate attachment sequences (attL2/attR2: AAAAATAATTAG). Once again, this finding suggests the possible role of incomplete excision of a previous prophage, followed by insertion of a new one, creating this composite phage in M92. ϕSa3 in strains TW20 and CMRSA6 are nearly identical, differing primarily in two hypothetical proteins. ϕSa3 of CMRSA3 is also very similar to the corresponding phage in TW20 and CMRSA6, once again differing primarily in hypothetical proteins, as well as in some of the DNA metabolism related proteins (Supplementary Table 2). ϕSa3 of M92, in contrast, is significantly different than ϕSa3 of the other strains. While it shares homology in the integrase and some tail, lysis and virulence associate proteins, it differs significantly in most of the proteins associated with DNA metabolism, portal and head proteins, as well as with some of the tail associated proteins

TABLE 1 | Antibiotic resistance profiles.

	Ampicillin	Amp/sul	β-lactam	Cefazolin	Penicillin	Oxacillin	Gentamicin	Erythromycin	Tetracycline	Ciprofloxacin	Clindamycin	Trimethoprim	TMP-SMX	Rifampin	Vancomycin	Nitrofurantoin	Neomycin
CMRSA6	R	R	P	R	R	R	R	R	R	R	R		R	S	S	S	
CMRSA3	R	R	P	R	R	R	S	R	R	R	R		R	S	S	S	
M92	R	R	P	R	R	R	R	R	R	R	R		S	S	S	S	
TW20	R	R			R	R	R	R	R	R		R					R

R, resistant; P, positive; S, sensitive; Amp/sul, ampicillin/sulbactam; TMX-SMX, trimethoprim/sulfamethoxazole.

(**Figure 6B**). The genes corresponding to virulence in φSa3 include staphylokinase, enterotoxin A, chemotaxis-inhibiting protein (CHIPS), and staphylococcal complement inhibitor (SCIN).

Staphylokinase (SATW20_19380) is present in all four strains, but the gene in CMRSA3 only shares 97% sequence homology with the genes in the other strains. Staphylokinase interacts with host proteins, including alpha-defensins (bactericidal peptides of human neutrophils) and plasminogen (Bokarewa et al., 2006). Binding of staphylokinase to alpha-defensins abolishes their activity, thereby protecting the bacteria from the human innate immune system. Interaction of staphylokinase with plasminogen, on the other hand, forms active plasmin, a proteolytic enzyme that enables bacterial penetration into the surrounding tissues (Bokarewa et al., 2006). Enterotoxin A (SATW20_19410), by contrast, is present in TW20 and CMRSA6, but absent in CMRSA3 and M92. Enterotoxins, including enterotoxin A, are notable virulence factors associated with *S. aureus* and have been implicated in toxic-shock-like syndrome and food poisoning, as well as acting as super-antigens that stimulate T-cell proliferation (Ortega et al., 2010). Finally, the gene for SCIN is present in φSa3 of all four strains (SATW20_19360 of TW20) while the gene for CHIPS is only found in CMRSA3. SCIN of M92 differs slightly from the other strains in that it contains a Leu80Gln substitution, but the role of this substitution in terms of protein function is unknown. SCIN inhibits central complement convertase, which reduces phagocytosis of the opsonized organism, blocking all downstream effector functions (Rooijakkers et al., 2007). CHIPS binds to the receptors for C5a and *N*-formyl peptides, reducing leukocyte recruitment (de Haas et al., 2004) and reducing bacterial killing.

A third mobile genetic element found in the genome of these ST239 isolates is the φSPβ-like prophage. It is 127.2 kb and integrated near 2200 kb in the genome (**Figure 6A**). φSPβ-like was detected in TW20, CMRSA6 and M92, but is absent from CMRSA3. The attL/R sequences (ATTATTATAATT) were identified at both ends of the prophage, including two attL sites located in close proximity to each other (**Figure 6B** and Supplementary Table 3). The phage in all three strains is nearly identical, with very minor variations in a few hypothetical proteins. φSPβ-like prophage is a large phage and exhibits similarity with the φSPβ-like region of *S. epidermidis* RP62a (Gill et al., 2005). It contains genes associated with aminoglycoside resistance, making the treatment of infection difficult (Holden et al., 2010). This phage also contains genes which may have a role in persistence of the organism in hospital settings (Holden et al., 2010). Since TW20, CMRSA6, and M92 belong to ST239, a major hospital associated MLST type, this phage likely plays an important role in their maintenance in the strong antibiotic selective pressures found in hospital environments. The φSPβ-like phage does not show similarity with any other *S. aureus* prophage and has been shown to be present only in epidemic ST239 strains (Xia and Wolz, 2014). CMRSA3, although closely related, belongs to ST241, which could explain why it is devoid of this prophage. This would, in turn, possibly explain why CMRSA3 (originally one of 10 epidemic MRSA strains) had virtually disappeared in Canada after 1997, being

FIGURE 6 | Six regions containing phages or mobile genetic elements are found in the ST239 isolates, with variations in their content. **(A)** BRIG alignment of the ST239 strains, with the locations of each MGE indicated. Rings from inner to outer (and darkest blue to lightest blue) are as follows: TW20, CMRSA6, CMRSA3, M92. **(B)** Comparisons of each mobile element region, with components differentiated by color. Homologies are marked with red for normal and blue for inverted, with color saturation indicative of percentage homology. Left and right MGE attachment site locations are indicated above the alignments, with markings below a strain representing differences within that strain. Attachment site sequences are as follows: φSa6/5 attL/R-AAAAAAGGGCAGA, attR2/L2-CTTTTTAAAATTA, attR3/L3-TAATTTAGTTAT; φSa3 attL/R-AAGTTGCAACAC, attL2/R2-AAAAATAATTAG; φSPβ-like attL/R- ATTATTATAATT; SaPI1 attL/R- TTGAAAATAAAA. Strains not depicted are missing the MGE in question. †While φSa6 of TW20, CMRSA6, and CMRSA3 are found in different genomic locations than φSa5 of M92, the phages are similar and included in the same comparison. *The φSPβ-like phage proteins in TW20, CMRSA6, and M92 are compared to φSPβ-like.

TABLE 2 | Summary of the genetic variation among the ST239 sub-lineage strains.

Genomic component	ST239 strain			
	TW20	**CMRSA6**	**CMRSA3**	**M92**
spa	+	+	+	−
lpl	Intact	Disrupted	Disrupted	Disrupted
cta	Intact	IS256 insertion	Intact	Intact
dfrG[†]	+	+	+ (two copies)	−
gyrA[†]	Ser84Leu	Ser84Leu	Ser84Leu	Ser84Leu
Isoleucyl-tRNA synthetase[†]	Val588Phe	Val	Isl581Phe	Val
φSa6[†] - No proteins related to virulence	+	+*	+**	−
φSa5 - No proteins related to virulence	−	−	−	+***
φSa3[†]	+	+*	+*	+***
-Staphylokinase	+	+	+	+
-Enterotoxin A	+	+	−	−
-SCIN	+	+	+	+
-CHIPS	−	−	+	−
φSPβ-like[†] - Resistance and persistence genes	+	+*	−	+*
- *sasX* (LPxTG protein)[†]	+	+	−	+
φSPβ-like proteins - No proteins related to virulence	+	+*	−	+*
SaPI1[†]	+	+*	+*	+*
-Enterotoxin K	+	+	+	+
-Enterotoxin Q	+	+	+	+
Major plasmid[†]	pTW20_1	−	pCMRSA3_1	−
-QacA	+		+	
-Mercury resistance	+		+	
-Cadmium resistance	+		+	
Minor plasmid[†] - No proteins related to virulence	pTW20_2	−	−	−

[†]*Loci/elements proposed by Holden et al. (2010) to contribute to the virulence of TW20. * Nearly identical to the region in TW20 (≥99% homology). ** Very similar to the region in TW20, with minor variations in protein content (90% homology). *** Significantly different from the region in TW20, with small regions of homology (≤50% overall homology).*

replaced by another closely related epidemic strain CMRSA6 (Christianson et al., 2007). In addition to φSPβ-like prophage, the genomes of TW20, CMRSA6, and M92 also carry φSPβ-like proteins near 1400 kb on their genomes (**Figure 6A**). These proteins show similarity to proteins in the φSPβ-like prophage (**Figure 6B** and Supplementary Table 4) and likely represent prophage remnants. The φSPβ-like proteins do not appear to contribute to virulence of the strains as no genes related to virulence were detected.

The final mobile genetic element detected in the genomes of TW20, CMRSA6, CMRSA3 and M92, is a pathogenicity island, SaPI1. It is located near 960 kb and contains attL/attR sequences (TTGAAAATAAAA) on each end (**Figures 6A,B** and Supplementary Table 5). The pathogenicity island proteins are nearly identical in all four strains. SaPI1 contains genes coding for enterotoxins K and Q (SATW20_08900 and SATW20_08910 of TW20 respectively), which are reported to play an important role in staphylococcal diseases (like food poisoning), as mentioned earlier.

Plasmids Identified in TW20 and CMRSA3

Extrachromosomal genetic material often carries resistance determinants and genes essential for virulence and pathogenicity

of an organism. Among TW20, CMRSA6, CMRSA3 and M92, only TW20 and CMRSA3 were found to carry plasmids. The plasmids of TW20, TW20_1, and TW20_2 are 29.5 and 3 kb respectively (Holden et al., 2010). TW20_1 carries important resistance determinants, including the gene coding for QacA, an antiseptic resistance protein conferring resistance to quaternary ammonium salts, cationic biocides and diamidines. TW20_1 also carries the *mer* and *cad* operons, containing genes coding for resistance to mercury (MerA, mercuric reductase) and cadmium (CadA, a cadmium transporting ATPase and CadD, cadmium resistance protein), respectively. The second plasmid carried by TW20, TW20_2, is approximately 3 kb and codes only for a replication origin and hypothetical protein. This plasmid is unlikely to contribute to virulence of the strain.

The plasmid, pCMRSA3 carried by CMRSA3 is 27 kb long and has components resembling pTW20_1 of TW20 and pZ172 of *S. aureus* subsp. *aureus* Z712. The arrangement of genes within these plasmids varies, however, the complementary regions among them share 99% homology. Like pTW20_1 and pZ172, pCMRSA3 carries genes for antiseptic resistance (*qacA*), mercury resistance (*merA*), and cadmium resistance (*cadD* and *cadA*).

Biocides and quaternary ammonium compounds are used as antiseptics on body surfaces, as well as disinfectants on

equipment and surfaces in many environments such as hospitals or farms. These compounds are also being used to improve hygiene, while some of the heavy metals that are relatively non-toxic to mammalian tissue are used as antimicrobial coatings and wound dressings (Wales and Davies, 2015). Resistance to these agents provides survival benefits to the organisms.

Antibiotic Resistance Profiles and Other Genes Contributing to Virulence

Resistance to multiple types of antibiotics plays an important role in the ability of an organism to survive, particularly within hospital environments where there is significant antibiotic selective pressure. The antibiotic resistance profiles of CMRSA6, CMRSA3, and M92 reveal that they are resistant to the majority of the antibiotics used in hospitals, a characteristic typical of HA-MRSA strains (**Table 1**). Their resistance profiles are nearly identical, with the exception of trimethoprim/sulfamethoxazole (where CMRSA6 and CMRSA3 are resistant, while M92 is susceptible) and gentamicin (where CMRSA6 and M92 are resistant, while CMRSA3 is susceptible). As discussed earlier, M92 lacks *dfrG*, which codes for the trimethoprim-resistant dihydrofolate reductase, making it susceptible to trimethoprim (**Figure 5A** and **Table 1**). CMRSA6 and CMRSA3, in contrast, both have the *dfrG* gene, protecting them from the action of trimethoprim (**Figure 5** and **Table 1**). Since limited data is available from the publication regarding the resistance profile of TW20 (Holden et al., 2010), we could not compare it fully to CMRSA6, CMRSA3, and M92. However, the available information does indicate that TW20's resistance pattern is consistent with that of a HA-MRSA, as it is resistant to the core antibiotics used in hospitals (including trimethoprim and gentamicin), in addition to β-lactams.

Surface anchored proteins with LPxTG motif bind host molecules and have been shown to be present in only 7% of ST239 strains. The presence of a LPxTG motif surface-anchored protein in TW20 (*sasX*; SATW20_21850) is proposed to be linked to its increased virulence and invasive capacity (Holden et al., 2010). Interestingly, this protein is also detected in the genomes of moderately virulent CMRSA6 and avirulent M92, suggesting that it likely plays a minor role in the augmented virulence of TW20. The gene was not detected in CMRSA3, which belongs to ST241, correlating well with findings that orthologs of this protein have not been detected in the sequenced genomes of *S. aureus* other than ST239 (Holden et al., 2010).

DNA gyrase subunit A of TW20 has a point mutation resulting in the substitution of Ser to Leu at position 84, which is in contrast to the majority of *S. aureus* strains that contain serine at that position. Studies have demonstrated that the Ser84Leu substitution is associated with resistance to quinolones, which may promote survival in the hospital environment. This point mutation was detected in all four of the ST239 isolates analyzed here.

Isoleucyl-tRNA synthetase in TW20 contains a Val588Phe substitution, which has been shown to confer chromosomal low-level mupirocin resistance. This substitution was not detected

in CMRSA6, CMRSA3, or M92. CMRSA3 did, however, have an isoleucine instead of phenylalanine at position 581, the significance of which is unknown.

CONCLUSION

Comparative genomic analysis of TW20, CMRSA6, CMRSA3, and M92 reveals remarkable similarities. Pulsotypes and SNP WGS phylogenetic cladograms of all four strains show that they cluster together forming a genomicly closely-related ST239 sub-lineage. While TW20 is positive for every genetic trait put forth as a possible contributor to virulence, CMRSA6, CMRSA3, and M92 showed variations in terms of carrying these traits (**Table 2**).

The major components differing among these strains are staphylococcal protein A (SpA) and the lipoprotein-like proteins (Lpl). The SpA of TW20, CMRSA6 and CMRSA3 is identical, while M92 is likely devoid of a functional *spa* gene encoded protein (**Table 2**). Similarly, the Lpl transcribed by TW20 has recently been shown to play a vital role in the pathogenesis of *S. aureus* species, but is disrupted in CMRSA6, CMRSA3, and M92. Mobile genetic elements may also play a role in the virulence of these strains. Three prophages (φSa6, φSa3, and φSPβ-like), a pathogenicity island (SaPI1) and two plasmids were located in strain TW20, and are present with variations in one or more of the other three strains (**Table 2**). Despite similar mobile element carriage, it is important to highlight the fact that even if similar virulence factors are present, these mobile elements are not identical; they have slight variations in their content with respect to hypothetical proteins of unknown function, any of which could play a significant role in the pathogenesis of the strain.

Further studies are needed to examine each of the genomic components brought forth in this study, with the goal of determining their exact contribution to the virulence and pathogenesis of TW20, CMRSA6, CMRSA3, and M92. None of these components exists in isolation, meaning the full virulence of *S. aureus* ST239 likely results from the sum of, and interplay between multiple factors.

AUTHOR CONTRIBUTIONS

KZ conceived, designed, and supervised the work. J-AM and AK performed the experiments and analyzed data. JC provided the clinical information. SL and KZ structured and drafted the manuscript. J-AM, JC, and KZ reviewed and edited the manuscript.

FUNDING

for Antimicrobial Resistance (CAR), Alberta Health Services, Calgary, AB, Canada.

ACKNOWLEDGMENTS

We thank T. Louie (University of Calgary, Canada) for generously providing us with the strain M92, J. Parkhill (The Wellcome Trust Sanger Institute, United Kingdom) for the strain

TW20, and S. Shideler (University of Calgary, Canada) for assistance with whole genome assembly.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01531/full#supplementary-material

REFERENCES

Aires de Sousa, M., and de Lencastre, H. (2004). Bridges from hospitals to the laboratory: genetic portraits of methicillin-resistant *Staphylococcus aureus* clones. *FEMS Immunol. Med. Microbiol.* 40, 101–111.doi: 10.1016/S0928-8244(03)00370-5

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402

Arndt, D., Grant, J., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387

Baba, T., Bae, T., Schneewind, O., Takeuchi, F., and Hiramatsu, K. (2008). Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. *J. Bacteriol.* 190, 300–310. doi: 10.1128/JB.01000-07

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., and Aravind, L., et al. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J. Bacteriol.* 188, 2761–2773. doi: 10.1128/JB.188.8.2761-2773.2006

Bokarewa, M. I., Jin, T., and Tarkowski, A. (2006). *Staphylococcus aureus*: Staphylokinase. *Int. J. Biochem. Cell Biol.* 38, 504–509. doi: 10.1016/j.biocel.2005.07.005

Christianson, S., Golding, G. R., Campbell, J., Canadian Nosocomial Infection Surveillance Program and Mulvey, M. R. (2007). Comparative genomics of Canadian epidemic lineages of methicillin-resistant *Staphylococcus aureus*. *J. Clin. Microbiol.* 45, 1904–1911. doi: 10.1128/JCM.02500-06

Clements, M. O., Watson, S. P., Poole, R. K., and Foster, S. J. (1999). CtaA of *Staphylococcus aureus* is required for starvation survival, recovery, and cytochrome biosynthesis. *J. Bacteriol.* 181, 501–507.

Conceicao, T., Aires-de-Sousa, M., Fuzi, M., Toth, A., Paszti, J., Ungvari, E., et al. (2007). Replacement of methicillin-resistant *Staphylococcus aureus* clones in Hungary over time: a 10-year surveillance study. *Clin. Microbiol. Infect.* 13, 971–979. doi: 10.1111/j.1469-0691.2007.01794.x

David, M. Z., and Daum, R. S. (2010). Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. *Clin. Microbiol. Rev.* 23, 616–687. doi: 10.1128/CMR.00081-09

de Haas, C. J., Veldkamp, K. E., Peschel, A., Weerkamp, F., Heezius, E. C., Poppelier, M. J., et al. (2004). Chemotaxis inhibitory protein of *Staphylococcus aureus*, a bacterial antiinflammatory agent. *J. Exp. Med.* 199, 687–695. doi: 10.1084/jem.20031636

de Vries, L. E., Hasman, H., Jurado Rabadan, S., and Agerso, Y. (2016). Sequence-based characterization of Tn5801-like genomic Islands in tetracycline-resistant *Staphylococcus pseudintermedius* and other gram-positive bacteria from humans and animals. *Front. Microbiol.* 7:576. doi: 10.3389/fmicb.2016.00576

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., et al. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13–W17. doi: 10.1093/nar/gkr245

Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J., and Spratt, B. G. (2000). Multilocus sequence typing for characterization of methicillin-resistant and

methicillin-susceptible clones of *Staphylococcus aureus*. *J. Clin. Microbiol.* 38, 1008–1015.

Enright, M. C., Robinson, D. A., Randle, G., Feil, E. J., Grundmann, H., and Spratt, B. G. (2002). The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc. Natl. Acad. Sci. U.S.A.* 99, 7687–7692. doi: 10.1073/pnas.122108599

Forsgren, A. (1970). Significance of protein a production by staphylococci. *Infect. Immun.* 2, 672–673.

Foster, T. J. (2005). Immune evasion by staphylococci. *Nat. Rev. Microbiol.* 3, 948–958. doi: 10.1038/nrmicro1289

Gill, S. R., Fouts, D. E., Archer, G. L., Mongodin, E. F., Deboy, R. T., Ravel, J., et al. (2005). Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J. Bacteriol.* 187, 2426–2438. doi: 10.1128/JB.187.7.2426-2438.2005

Goodyear, C. S., and Silverman, G. J. (2004). Staphylococcal toxin induced preferential and prolonged in vivo deletion of innate-like B lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11392–11397. doi: 10.1073/pnas.0404382101

Gray, R. R., Tatem, A. J., Johnson, J. A., Alekseyenko, A. V., Pybus, O. G., Suchard, M. A., et al. (2011). Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a bayesian framework. *Mol. Biol. Evol.* 28, 1593–1603. doi: 10.1093/molbev/msq319

Hammer, N. D., Reniere, M. L., Cassat, J. E., Zhang, Y., Hirsch, A. O., Indriati Hood, M., et al. (2013). Two heme-dependent terminal oxidases power *Staphylococcus aureus* organ-specific colonization of the vertebrate host. *MBio* 4:e00241-13. doi: 10.1128/mBio.00241-13

Hammer, N. D., Schurig-Briccio, L. A., Gerdes, S. Y., Gennis R. B., and Skaar, E. P. (2016). CtaM is required for menaquinol oxidase aa3 function in *Staphylococcus aureus*. *MBio* 7:e00823-16. doi: 10.1128/mBio.00823-16

Hanson, D. C., and Schumaker, V. N. (1984). A model for the formation and interconversion of protein A-immunoglobulin G soluble complexes. *J. Immunol.* 132, 1397–1409.

Harmsen, D., Claus, H., Witte, W., Rothganger, J., Claus, H., Turnwald, D., et al. (2003). Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for spa repeat determination and database management. *J. Clin. Microbiol.* 41, 5442–5448. doi: 10.1128/JCM.41.12.5442-5448.2003

Harris, S. R., Feil, E. J., Holden, M. T., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395

Hjelm, H., Sjodahl J., and Sjoquist, J. (1975). Immunologically active and structurally similar fragments of protein A from *Staphylococcus aureus*. *Eur. J. Biochem.* 57, 395–403. doi: 10.1111/j.1432-1033.1975.tb02313.x

Holden, M. T., Lindsay, J. A., Corton, C., Quail, M. A., Cockfield, J. D., Pathak, S., et al. (2010). Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). *J. Bacteriol.* 192, 888–892. doi: 10.1128/JB.01255-09

Hsu, L. Y., Harris, S. R., Chlebowicz, M. A., Lindsay, J. A., Koh, T. H., Krishnan, P., et al. (2015). Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol.* 16:81. doi: 10.1186/s13059-015-0643-z

Hyman, P., Abedon, S. T., and Centre for Agriculture and Bioscience International (2012). *Bacteriophages in Health and Disease*. Wallingford: CABI. doi: 10.1079/9781845939847.0000

Kahankova, J., Pantucek, R., Goerke, C., Ruzickova, V., Holochova, P., and Doskar, J. (2010). Multilocus PCR typing strategy for differentiation of *Staphylococcus aureus* siphoviruses reflecting their modular genome structure. *Environ. Microbiol.* 12, 2527–2538. doi: 10.1111/j.1462-2920.2010.02226.x

Kock, R., Becker, K., Cookson, B., van Gemert-Pijnen, J. E., Harbarth, S., Kluytmans, J., et al. (2010). Methicillin-resistant *Staphylococcus aureus* (MRSA): burden of disease and control challenges in Europe. *Euro Surveill.* 15:19688. doi: 10.2807/ese.15.41.19688-en

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., et al. (2001). Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* 357, 1225–1240. doi: 10.1016/S0140-6736(00)04403-2

Lan, L., Cheng, A., Dunman, P. M., Missiakas, D., and He, C. (2010). Golden pigment production and virulence gene expression are affected by metabolisms in *Staphylococcus aureus*. *J. Bacteriol.* 192, 3068–3077. doi: 10.1128/JB.00928-09

McClure, J. A., Conly, J. M., Elsayed, S., and Zhang, K. (2010). Multiplex PCR assay to facilitate identification of the recently described Staphylococcal cassette chromosome mec type VIII. *Mol. Cell. Probes* 24, 229–232. doi: 10.1016/j.mcp.2010.01.001

McClure, J. A., and Zhang, K. (2017). Complete genome sequence of the methicillin-resistant *Staphylococcus aureus* colonizing strain M92. *Genome Announc.* 5:e00478-17. doi: 10.1128/genomeA.00478-17

Miovic, M., and Pizer, L. I. (1971). Effect of trimethoprim on macromolecular synthesis in *Escherichia coli*. *J. Bacteriol.* 106, 856–862.

Mogi, T., Saiki, K., and Anraku, Y. (1994). Biosynthesis and functional role of haem O and haem A. *Mol. Microbiol.* 14, 391–398. doi: 10.1111/j.1365-2958.1994.tb02174.x

Monecke, S., Coombs, G., Shore, A. C., Coleman, D. C., Akpaka, P., Borg, M., et al. (2011). A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. *PLoS One* 64:e17936. doi: 10.1371/journal.pone.0017936

Moretti, S., Armougom, F., Wallace, I. M., Higgins, D. G., Jongeneel, C. V., and Notredame, C. (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res.* 35, W645–W648. doi: 10.1093/nar/gkm333

Mulvey, M. R., Chui, L., Ismail, J., Louie, L., Murphy, C., Chang, N., et al. (2001). Development of a Canadian standardized protocol for subtyping methicillin-resistant *Staphylococcus aureus* using pulsed-field gel electrophoresis. *J. Clin. Microbiol.* 39, 3481–3485. doi: 10.1128/JCM.39.10.3481-3485.2001

Nathwani, D., Morgan, M., Masterton, R. G., Dryden, M., Cookson, B. D., French, G., et al. (2008). Guidelines for UK practice for the diagnosis and management of methicillin-resistant *Staphylococcus aureus* (MRSA) infections presenting in the community. *J. Antimicrob. Chemother.* 61, 976–994. doi: 10.1093/jac/dkn096

Neoh, H. M., Cui, L., Yuzawa, H., Takeuchi, F., Matsuo, M., and Hiramatsu, K. (2008). Mutated response regulator graR is responsible for phenotypic conversion of *Staphylococcus aureus* from heterogeneous vancomycin-intermediate resistance to vancomycin-intermediate resistance. *Antimicrob. Agents Chemother.* 52, 45–53. doi: 10.1128/AAC.00534-07

Nguyen, M. T., Kraft, B., Yu, W., Demircioglu, D. D., Hertlein, T., Burian, M., et al. (2015). The vSaα Specific Lipoprotein Like Cluster (lpl) of *S. aureus* USA300 contributes to immune stimulation and invasion in human cells. *PLoS Pathog.* 11:e1004984. doi: 10.1371/journal.ppat.1004984

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205–217. doi: 10.1006/jmbi.2000.4042

Ortega, E., Abriouel, H., Lucas, R., and Galvez, A. (2010). Multiple roles of *Staphylococcus aureus* enterotoxins: pathogenicity, superantigenic activity, and correlation to antibiotic resistance. *Toxins* 2, 2117–2131. doi: 10.3390/toxins2082117

Peacock, S. J., Moore, C. E., Justice, A., Kantzanou, M., Story L., Mackie, K., et al. (2002). Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. *Infect. Immun.* 70, 4987–4996. doi: 10.1128/IAI.70.9.4987-4996.2002

Rooijakkers, S. H., Milder, F. J., Bardoel, B. W., Ruyken, M., van Strijp, J. A., and Gros, P. (2007). Staphylococcal complement inhibitor: structure and active sites. *J. Immunol.* 179, 2989–2998. doi: 10.4049/jimmunol.179.5.2989

Rooijakkers, S. H. M., van Kessel, K. P. M., and van Strijp, J. A. G. (2005). Staphylococcal innate immune evasion. *Trends Microbiol.* 13, 596–601. doi: 10.1016/j.tim.2005.10.002

Rouch, D. A., Messerotti, L. J., Loo, L. S., Jackson, C. A., and Skurray, R. A. (1989). Trimethoprim resistance transposon Tn4003 from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS257. *Mol. Microbiol.* 3, 161–175. doi: 10.1111/j.1365-2958.1989.tb01805.x

Schneewind, O., Mihaylova-Petkov, D., and Model, P. (1993). Cell wall sorting signals in surface proteins of gram-positive bacteria. *EMBO J.* 12, 4803–4811.

Schneewind, O., Model, P., and Fischetti, V. A. (1992). Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267–281. doi: 10.1016/0092-8674(92)90101-H

Shahmirzadi, S. V., Nguyen, M. T., and Gotz, F. (2016). Evaluation of *Staphylococcus aureus* lipoproteins: role in nutritional acquisition and pathogenicity. *Front. Microbiol.* 7:1404. doi: 10.3389/fmicb.2016.01404

Sjodahl, J. (1977a). Repetitive sequences in protein A from *Staphylococcus aureus*. Arrangement of five regions within the protein, four being highly homologous and Fc-binding. *Eur. J. Biochem.* 73, 343–351. doi: 10.1111/j.1432-1033.1977.tb11324.x

Sjodahl, J. (1977b). Structural studies on the four repetitive Fc-binding regions in protein A from *Staphylococcus aureus*. *Eur. J. Biochem.* 78, 471–490. doi: 10.1111/j.1432-1033.1977.tb11760.x

Stevens, D. L., Bisno, A. L., Chambers, H. F., Dellinger, E. P., Goldstein, E. J. C., Gorbach, S. L., et al. (2014). Executive summary: practice guidelines for the diagnosis and management of skin and soft tissue infections: 2014 update by the infectious diseases society of America. *Clin. Infect. Dis.* 59, 147–159. doi: 10.1093/cid/ciu444

Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039

Svensson, B., and Hederstedt, L. (1994). Bacillus subtilis CtaA is a heme-containing membrane protein involved in heme A biosynthesis. *J. Bacteriol.* 176, 6663–6671. doi: 10.1128/jb.176.21.6663-6671.1994

Svensson, B., Lubben, M., and Hederstedt, L. (1993). Bacillus subtilis CtaA and CtaB function in haem A biosynthesis. *Mol. Microbiol.* 10, 193–201. doi: 10.1111/j.1365-2958.1993.tb00915.x

Tsuru, T., and Kobayashi, I. (2008). Multiple genome comparison within a bacterial species reveals a unit of evolution spanning two adjacent genes in a tandem paralog cluster. *Mol. Biol. Evol.* 25, 2457–2473. doi: 10.1093/molbev/msn192

Uhlen, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., and Lindberg, M. (1984). Complete sequence of the staphylococcal gene encoding protein A. A gene evolved through multiple duplications. *J. Biol. Chem.* 259, 1695–1702.

Wales, A. D., and Davies, R. H. (2015). Co-Selection of resistance to antibiotics, biocides and heavy metals, and its relevance to foodborne pathogens. *Antibiotics* 4, 567–604. doi: 10.3390/antibiotics4040567

Wallace, I. M., O'Sullivan, O., Higgins, D. G., and Notredame, C. (2006). M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* 34, 1692–1699. doi: 10.1093/nar/gkl091

Wang, S., Khan, Y., Hines, L., Mediavilla, J. R., Zhang, L., Chen, L., et al. (2012). Methicillin-resistant *Staphylococcus aureus* sequence type 239-III, Ohio, USA, 2007-2009. *Emerg. Infect. Dis.* 18, 1557–1565. doi: 10.3201/eid1810.120468

Wright, C., Willan, K. J., Sjodahl, J., Burton, D. R., and Dwek, R. A. (1977). The interaction of protein A and Fc fragment of rabbit immunoglobulin G as probed by complement-fixation and nuclear-magnetic-resonance studies. *Biochem. J.* 167, 661–668. doi: 10.1042/bj1670661

Wu, K., Conly, J., McClure, J. A., Elsayed, S., Louie, T., and Zhang, K. (2010). *Caenorhabditis elegans* as a host model for community-associated methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 16, 245–254. doi: 10.1111/j.1469-0691.2009.02765.x

Wu, K., Conly, J., Surette, M., Sibley, C., Elsayed, S., and Zhang, K. (2012a). Assessment of virulence diversity of methicillin-resistant *Staphylococcus aureus*

strains with a *Drosophila melanogaster* infection model. *BMC Microbiol.* 12:274. doi: 10.1186/1471-2180-12-274

Wu, K., Simor, A. E., Vearncombe, M., McClure, J. A., and Zhang, K. (2012b). A *Caenorhabditis elegans* host model correlates with invasive disease caused by *Staphylococcus aureus* recovered during an outbreak in neonatal intensive care. *Can. J. Infect. Dis. Med. Microbiol.* 23, 130–134. doi: 10.1155/2012/543817

Xia, G., and Wolz, C. (2014). Phages of *Staphylococcus aureus* and their impact on host evolution. *Infect. Genet. Evol.* 21, 593–601. doi: 10.1016/j.meegid.2013.04.022

Xu, T., Han, J., Zhang, J., Chen, J., Wu, N., Zhang, W., et al. (2016). Absence of protoheme IX farnesyltransferase CtaB causes virulence attenuation but enhances pigment production and persister survival in MRSA. *Front. Microbiol.* 7:1625. doi: 10.3389/fmicb.2016.01625

Zhang, K., McClure, J. A., and Conly, J. M. (2012). Enhanced multiplex PCR assay for typing of staphylococcal cassette chromosome mec types I to V in methicillin-resistant *Staphylococcus aureus*. *Mol. Cell. Probes* 26, 218–221. doi: 10.1016/j.mcp.2012.04.002

Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer BM and handling Editor declared their shared affiliation.

Significant Enrichment and Diversity of the Staphylococcal Arginine Catabolic Mobile Element ACME in *Staphylococcus epidermidis* Isolates From Subgingival Peri-implantitis Sites and Periodontal Pockets

Aoife M. O'Connor[1†], Brenda A. McManus[1†], Peter M. Kinnevey[1], Gráinne I. Brennan[2], Tanya E. Fleming[2], Phillipa J. Cashin[1], Michael O'Sullivan[3], Ioannis Polyzois[3] and David C. Coleman[1]*

[1] Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin, Ireland, [2] National MRSA Reference Laboratory, St. James's Hospital, Dublin, Ireland, [3] Division of Restorative Dentistry and Periodontology, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin, Ireland

Staphylococcus aureus and *Staphylococcus epidermidis* are frequent commensals of the nares and skin and are considered transient oral residents. Reports on their prevalence in the oral cavity, periodontal pockets and subgingivally around infected oral implants are conflicting, largely due to methodological limitations. The prevalence of these species in the oral cavities, periodontal pockets and subgingival sites of orally healthy individuals with/without implants and in patients with periodontal disease or infected implants (peri-implantitis) was investigated using selective chromogenic agar and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Staphylococcus epidermidis* was predominant in all participant groups investigated. Its prevalence was significantly higher ($P = 0.0189$) in periodontal pockets (30%) than subgingival sites of healthy individuals (7.8%), and in subgingival peri-implantitis sites (51.7%) versus subgingival sites around non-infected implants (16.1%) ($P = 0.0057$). In contrast, *S. aureus* was recovered from subgingival sites of 0-12.9% of the participant groups, but not from periodontal pockets. The arginine catabolic mobile element (ACME), thought to enhance colonization and survival of *S. aureus*, was detected in 100/179 *S. epidermidis* and 0/83 *S. aureus* isolates screened using multiplex PCR and DNA microarray profiling. Five distinct ACME types, including the recently described types IV and V (I; 14, II; 60, III; 10, IV; 15, V; 1) were identified. ACME-positive *S. epidermidis* were significantly ($P = 0.0369$) more prevalent in subgingival peri-implantitis sites (37.9%) than subgingival sites around non-infected implants (12.9%) and also in periodontal pockets (25%) compared to subgingival sites of healthy individuals (4.7%) ($P = 0.0167$). To investigate the genetic diversity of ACME, 35 isolates, representative of patient groups, sample sites and ACME types underwent whole genome sequencing from which multilocus sequence types (STs) were identified.

Sequencing data permitted ACME types II and IV to be subdivided into subtypes IIa-c and IVa-b, respectively, based on distinct flanking direct repeat sequences. Distinct ACME types were commonly associated with specific STs, rather than health/disease states or recovery sites, suggesting that ACME types/subtypes originated amongst specific *S. epidermidis* lineages. Ninety of the ACME-positive isolates encoded the ACME-*arc* operon, which likely contributes to oral *S. epidermidis* survival in the nutrient poor, semi-anaerobic, acidic and inflammatory conditions present in periodontal disease and peri-implantitis.

Keywords: ACME, *Staphylococcus epidermidis*, periodontal disease, peri-implantitis, subgingival sites, oral cavity, periodontal pockets, *kdp* operon

INTRODUCTION

Staphylococcus aureus and *Staphylococcus epidermidis* are common commensals of human skin and the nares and are highly proficient at forming biofilms. Both species are significant causes of nosocomial infections associated with indwelling medical devices (Song et al., 2013). It is now widely acknowledged that many of the antimicrobial resistance genes identified in clinical isolates of *S. aureus* were acquired from coagulase negative staphylococci (CoNS) such as *S. epidermidis* by transfer of mobile genetic elements (MGEs) (Otto, 2013). One very significant example of this is the horizontal acquisition by *S. aureus* of staphylococcal cassette chromosome-*mec* (SCC*mec*) elements harboring the methicillin resistance gene *mec* from *S. epidermidis*. To date, 13 different types of SCC*mec* (I-XIII) have been characterized in methicillin-resistant *S. aureus* (MRSA)[1] differentiated according to the various combinations of *mec* and cassette chromosome recombinase (*ccr*) gene complexes present (International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements, 2009; Wu et al., 2015; Baig et al., 2018). Many more SCC*mec* variants have been described in species in which methicillin-resistance (MR) is more common, such as *S. epidermidis* (MRSE) and other CoNS (Shore and Coleman, 2013).

The staphylococcal arginine catabolic mobile element (ACME) plays a role in colonization of human skin and evasion of host immune responses (Diep et al., 2008; Planet et al., 2013). It is considered an SCC-like element as, like SCC*mec*, it is flanked by homologous inverted and direct repeat sequences (DRs) that integrate into the same *attB* attachment site in the chromosomal *orfX* locus as SCC*mec* (Ito et al., 2001; Diep et al., 2006). Specific clonal lineages of *S. aureus* are known to harbor ACME, most notably the highly successful USA300 clone, which harbors a composite island (CI) composed of SCC*mec* type IVa and ACME type I (Diep et al., 2006). The prevalence of ACME has also been reported to be high in sequence type (ST) 239 MRSA isolates (43.7%) from screening swabs of hospitalized patients in Singapore (Hon et al., 2014) and in ST239-like (as determined by pulsed-field gel electrophoresis) bloodstream MRSA isolates (39%) recovered in Australia (Espedido et al., 2012).

Like SCC*mec*, ACME is more prevalent and exhibits greater diversity in *S. epidermidis*. Many studies have identified ACME in multiple STs of the predominant *S. epidermidis* clonal lineages based on multilocus sequence typing (MLST), suggesting that ACME originated in this species (Miragaia et al., 2009; Barbier et al., 2011; Onishi et al., 2013). To date, ACME has been detected in the range of 45.8–67.9% in *S. epidermidis* isolates recovered from disparate geographical locations, as well as in carriage and disease isolates (Diep et al., 2006; Miragaia et al., 2009; Barbier et al., 2011; Onishi et al., 2013).

The ACME genetic island ranges between 30 and 55 kb in size and is associated with three main gene clusters, the *arc* operon composed of the *arcR/A/D/B/C* genes, the *opp3* operon composed of the *opp3A/B/C/D/E* genes, and the recently revealed *kdp* operon, composed of the *kdpE/D/A/B/C* genes (Diep et al., 2006; O'Connor et al., 2018). These gene clusters encode an arginine deaminase pathway, an oligopeptide permease ABC transporter and a potassium ABC transporter, respectively. These three operons which can be present in ACME are in addition to the native chromosomal *arc*, *opp1* and *opp2*, and *kdp* operons in *S. aureus* (Diep et al., 2008; Xue et al., 2011; Price-Whelan et al., 2013).

Five distinct ACME types have been described to date, according to the presence of the *arc* and *opp3* operons (type I), the *arc* operon only (type II), the *opp3* operon only (type III), the *arc* and *kdp* operons (type IV), and all three *arc*, *opp* and *kdp* operons (type V) (Gill et al., 2005; Diep et al., 2006; Shore et al., 2011; McManus et al., 2017; O'Connor et al., 2018). Furthermore, two distinct ACME IV subtypes, IVa and IVb have been described based on distinct combinations of flanking DRs (O'Connor et al., 2018). To date, all five types and several variants thereof have been described in *S. epidermidis* (Miragaia et al., 2009; Barbier et al., 2011; Onishi et al., 2013; Soroush et al., 2016; McManus et al., 2017; O'Connor et al., 2018). In contrast, types I and II and variants thereof have been detected in *S. aureus*, commonly collocated with other genetic elements such as SCC*mec* or SCC-associated genes in CIs and separated from these adjacent elements by DRs (Diep et al., 2006; Shore et al., 2011; Kawaguchiya et al., 2013; Rolo et al., 2012).

Although staphylococci are considered transient members of the oral microflora, these species are prevalent in the oral cavities of the elderly and in people with dental infections such as periodontal disease (Murdoch et al., 2004; Friedlander,

---

[1]www.SCCmec.org

2010). Periodontal disease is an inflammatory condition that can progress from gingivitis in response to dental plaque and affects the gingiva as well as the supporting periodontal structures (Hajishengallis, 2015). As periodontal disease progresses, enlargement of the gingival crevice occurs and leads to eventual detachment of the gingival tissue from the tooth resulting in periodontal pocket formation. Periodontal pockets provide a semi-anaerobic nutrient-poor environment that is ideal for plaque accumulation by resident oral microflora and is prone to decreases in pH resulting from physiological processes such as tissue repair (Percival et al., 2014).

The titanium-based oral implant can act as an ideal substrate for staphylococcal-based biofilm formation (Thurnheer and Belibasakis, 2016), as can the oxygen-depleted environment of periodontal and peri-implantitis pockets. Dental implants are indwelling medical devices made of titanium-based alloys that are placed in the bone of the mandible or maxilla to anchor a prosthetic crown, denture or bridge (Adell, 1981; Branemark et al., 1983). They consist of a shaft that is placed directly in the jaw bone and stabilized by subsequent osseointegration, and an abutment onto which a prosthesis is fitted. Similar to gingivitis, peri-implant mucositis is an inflammatory condition that affects the gingivae surrounding a dental implant, which can progress to peri-implantitis in which supporting bone surrounding an implant is gradually lost, potentially resulting in implant failure (Renvert and Polyzois, 2018).

Both of these oral diseases have a similar etiology in that they are both associated with dental plaque in which a shift from normal resident microflora to more periodontopathogenic species appears to occur (Nibali et al., 2015).

This study investigated the prevalence of *S. epidermidis* and *S. aureus* in the oral cavities, subgingival sites and periodontal pockets of patients with implants and natural teeth in states of both health and disease. Isolates recovered were investigated for ACME to determine if ACME could be a molecular marker for periodontal disease and/or peri-implantitis. Previous studies investigated the prevalence of ACME in both *S. aureus* and *S. epidermidis* in a range of carriage and infection sites (Miragaia et al., 2009; Barbier et al., 2011; Du et al., 2013; Onishi et al., 2013), however, to our knowledge, no studies have investigated the prevalence of ACME among oral staphylococcal isolates from periodontal pockets or peri-implantitis sites. A selection of the ACME-positive isolates identified in the present study were further investigated by whole-genome sequencing (WGS) in order to elucidate the genetic organization of the ACMEs in detail. Such investigations could yield important information regarding the potential genetic reservoir of ACME that exist among *S. epidermidis* for potential future spread to MRSA.

MATERIALS AND METHODS

Study Group

Ethical approval for this study was granted by the St. James's Hospital and Federated Dublin Voluntary Hospitals Joint Research Ethics Committee (JREC) and the Faculty of Health Sciences Ethics Committee of Trinity College Dublin, Ireland. Prior to enrollment in the study, all participants were provided with comprehensive patient information documentation and all participants included provided written consent. All documentation (including consent forms) provided to patients was pre-approved by the Research Ethics Committees.

All participants in the study met the following criteria: they were over 18 years of age, had a minimum of 10 natural teeth and were capable of providing informed consent. Participants were excluded from the study if they had any of the following factors: diabetes or asthma, pregnancy or lactation, blood-borne illnesses, steroid treatment within the year or antibiotics within 2 months prior to sampling. Patients with periodontal disease had a minimum of one periodontal site with a probing depth of greater than 6 mm and bleeding on probing (BOP). Patients with peri-implantitis were partially dentate and had one or more oral implants in place for a minimum of 5 years, at least one of which showed clinical signs of disease (Sanz and Chapple, 2012; Holtfreter et al., 2015). The study group consisted of 31 orally healthy patients with dental implants, 20 patients with periodontal disease, 21 patients with peri-implantitis and 64 orally healthy participants.

Sample Collection and Processing

All clinical sampling was carried out by qualified Dentists at the Dublin Dental University Hospital (DDUH). Sub-gingival sites and periodontal pockets were sampled by inserting a PerioPaper[TM] gingival fluid collection strip (Oroflow, Plainview, NY, United States) into the sub-gingival crevice or periodontal pocket for 30 s. Following sampling the collection strips were placed in sterile 2 ml screw-capped tubes (Sarstedt AG & Co., Numbrecht, Germany) containing 1 ml of nutrient broth (NB) (Oxoid Ltd., Hampshire, United Kingdom). In addition, oral rinse samples were collected by providing participants with sterile 100 ml plastic cups (Sarstedt AG & Co.) containing 25 ml sterile phosphate buffered saline (PBS) and instructing the participant to rinse their mouths with the PBS for 30 s before returning the rinse fluid to the same container. The anterior nares of participants were sampled using nitrogen-gassed VI-packed sterile transport swabs (Sarstedt AG & Co.). Following sampling, all samples were transported immediately to the microbiology laboratory and processed within 4 h. Vials containing PerioPaper[TM] strips suspended in NB were vortexed at maximum speed for 1 min and 100 μl aliquots of the resulting cell suspension were plated onto mannitol salt agar (MSA) and Sa*Select*[TM] chromogenic agar (Bio-Rad Laboratories, Hertfordshire, United Kingdom) agar. Oral rinse samples were processed by transferring a 1 ml aliquot to a sterile 1.5 ml Eppendorf Safe-lock[TM] microfuge tube (Eppendorf, Hamburg, Germany) and centrifuged at $20,000 \times g$ for 1 min, after which the supernatant was discarded and the resultant pellet was resuspended in 200 μl NB. To isolate staphylococcal colonies, 100 μl aliquots of this cell suspension were plated on MSA and Sa*Select*[TM]. Nasal swabs were used to lawn the entire surface of MSA and Sa*Select*[TM] plates. Inoculated MSA and Sa*Select*[TM] plates were incubated at

37°C for 48 h in a static incubator (Gallenkamp, Leicester, United Kingdom).

Culture, Identification and Storage of Isolates

Bacterial isolates were cultured on Columbia blood agar (Fannin Ltd., Dublin, Republic of Ireland) at 37°C for 48 h prior to identification by Vitek MS Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry system (MALDI-TOF MS) (Vitek, bioMérieux Marcy l'Etoile, France) according to the manufacturer's instructions. Multiple isolates were identified and stored from each sample for future analysis. All isolates were stored on Microbank™ storage beads (Pro-Lab Diagnostics, Cheshire, United Kingdom) at −80°C.

DNA Isolation and Detection of ACME by Multiplex PCR and DNA Microarray Profiling

Where possible, a minimum of two isolates were selected as representatives of each individual participant, distinct sample sites and each staphylococcal species recovered and screened by multiplex PCR to detect the presence of ACME.

Genomic DNA was extracted from isolates by enzymatic lysis using the buffers and solutions provided with the *S. aureus* Genotyping Kit 2 DNA microarray kit (Alere Technologies GmbH, Jena, Germany) and the DNeasy Blood and Tissue kit (Qiagen, Crawley, West Sussex, United Kingdom) according to the manufacturers' instructions.

The presence of ACME was detected in isolates by multiplex PCR targeting the *arcA, opp3B* and *kdpA* genes harbored by ACME using the previously described *arcA*- and *opp3B*-directed primers (Diep et al., 2006; McManus et al., 2017) and incorporating primers targeting the *kdpA* gene (kdpF: 5′-CGGTTTAACTGGTGCGTT-3′ and kdpR: 5′-GCAATACATACAGCGTAGCC-3′) (O'Connor et al., 2018). PCR assays were carried out in 50 μl reaction volumes containing a 200 μM concentration of each deoxynucleoside triphosphate, 1.25 U of GoTaq polymerase (Promega, Madison, WI, United States), 10 μl (1×) of GoTaq FlexiBuffer (Promega), 2.5 μM MgCl$_2$, 100 pmol of each primer, and 1 ng of the DNA template. Cycling conditions consisted of 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 45 s and followed by a final elongation step of 72°C for 10 min. Amplification products (*arcA* product: 724 bp, *opp3B* product 530 bp, *kdpA* product: 241 bp) were separated by electrophoresis in 2% (w/v) agarose (Sigma-Aldrich Ltd., Wicklow, Republic of Ireland) containing 1X GelRed® (Biotium Inc., Fremont, CA, United States) and visualized using an Alpha Innotech UV transilluminator (Protein Simple, San Jose, CA, United States).

The presence of *mec* and ACME-*arc* genes amongst *S. aureus* and *S. epidermidis* isolates investigated was also detected by DNA microarray profiling using the *S. aureus* Genotyping Kit 2.0 (Alere Technologies GmbH, Jena, Germany) according to the manufacturer's instructions and as described previously (Monecke et al., 2008).

Molecular Characterization of ACME Elements by WGS

A total of 35 *S. epidermidis* isolates selected as representatives of each patient group, sample site and ACME type present were subjected to WGS (**Table 1**). Libraries were prepared using Nextera XT library preparation reagents (Illumina, Eindhoven, Netherlands) and sequenced using an Illumina MiSeq desktop sequencer. For each isolate, reads were aligned with reference *S. epidermidis* and *S. aureus* genomes containing ACME and/or SCC elements downloaded from Genbank using a Burrows-Wheeler aligner (BWA) (Li and Durbin, 2009) to select the most appropriate reference ACME type to use as a scaffold. *De novo* assemblies were carried out on the reads for each ACME-harboring isolate using SPAdes version 3.6[2]. For each isolate, the reference genome that exhibited the highest degree of alignment with the relevant reads was used in a further alignment with the annotated contigs from the *de novo* assembly of the relevant isolate. Contigs identified as containing ACME- or SCC*mec*-associated DNA sequences were aligned, annotated and visualized using BioNumerics version 7.6 (Applied Maths, Sint-Martens-Latem, Belgium) and the Artemis sequence viewer (Berriman and Rutherford, 2003).

In order to confirm the genetic organization and orientation of contigs, primers were designed using BioNumerics version 7.6 that targeted a minimum distance of 200 nucleotides from the contig boundaries. The target specificity of primers was confirmed using BLAST software[3]. All primers were supplied by Sigma–Aldrich Ltd. Contig gaps were closed based on PCR-based amplification and Sanger-based sequencing of these regions using the primers listed in Supplementary Table S1. Sanger-based sequencing was carried out commercially by Source BioScience (Waterford, Republic of Ireland).

Multiple alignments of the complete nucleotide sequences of the *arc*, *opp3* and *kdp* operons (including intragenic regions) from each isolate investigated were carried out using the Clustal Omega tool (Sievers et al., 2011). The nucleotide sequence from the first to last base of each operon, including any intragenic regions was compared among all isolates investigated in the present study.

Determination of STs Among Isolates Subjected to WGS

The STs of ACME-harboring *S. epidermidis* isolates subjected to WGS were determined from the WGS data by examination of the nucleotide sequences of the loci used for the consensus *S. epidermidis* MLST scheme (Thomas et al., 2007). Briefly, the relevant nucleotide sequences were gleaned from the WGS data and inputted into the *S. epidermidis* MLST database online[4] in order to define allelic profiles and STs.

---

[2]http://cab.spbu.ru/software/spades/
[3]https://blast.ncbi.nlm.nih.gov/Blast.cgi
[4]http://sepidermidis.mlst.net/

TABLE 1 | Population of *Staphylococcus epidermidis* isolates subjected to whole genome sequencing.

Isolate	Sample site	Patient	ACME type/subtype	ACME/CI size (kb)	ST[a]	Genbank accession number[b]
Participants with periodontal disease (*n* = 12)						
P8OR3	OR	P8	IVb	54.2	210	MG787414[b]
P9OR1	OR	P9	IIb	31.1	701	MH188462
P9PPH12	SG	P9	IIb	31.1	73	MH188463
P9PPHI1	SG	P9	IIb	31.1	73	MH188464
P11OR1	OR	P11	IIb	31.1	73	MH188465
P11PPH21	SG	P11	IIb	31.1	73	MH188466
P11PPP12	PP	P11	IIa	54.4	59	MH188467
P14OR1	OR	P14	I	54.3	17	MH188468
P14PPP2	PP	P14	IIc	53.9	672	MH188469
P14NS2	NS	P14	IIb	30.7	14	MH188470
P16OR1	OR	P16	III	53.9	329	MF346684[b]
P19PPP1	PP	P19	IIc	53.5	672	MH188471
Participants with peri-implantitis (*n* = 9)						
PS36PD	SG	PS36	IVa	40.2	432	MG787422[b]
PS7OR	OR	PS7	IIb	32	73	MH188472
PS7P2	SG	PS7	IVa	67.8	153	MH188473
PS23P1	SG	PS23	IVb	54.2	153	MH188474
PS30PH	SG	PS30	IVa	68.3	153	MG787421[b]
PS34PI	SG	PS34	IIb	31.6	14	MH188475
PS8TI	SG	PS8	IIb	32	73	MH188476
PS19PH	SG	PS19	V	116.9	5	MG787423[b]
PS21NS	NS	PS21	IVa	55.8	297	MG787420[b]
Orally healthy participants with implants (*n* = 6)						
I9OR1	OR	I9	IVa	67.8	153	MG787415[b]
I11OR1	OR	I11	III	45.1	329	MF346685[b]
I12OR1	OR	I12	I	39.9	7	MH188477
I14OR1	OR	I14	IVa	67.8	153	MG787416[b]
I23OR2	OR	I23	IIb	48	89	MH188478
I1PPP121	SG	I1	III	66.9	329	MH188479
Orally healthy participants (*n* = 9)						
120PPC	SG	120	IVa	67.8	153	MG787417[b]
200OR2	OR	200	IIb	32	73	MH188480
201OR2	OR	201	IIa	27	59	MH188481
204OR1	OR	204	III	65.6	329	MF346683[b]
217PPP362	SG	217	IIa	74.6	59	MH188482
218PP361	SG	218	IVa	39	130	MG787418[b]
32BR	OR	32	IIb	32	73	MH188483
33BR	OR	33	IVa	48.8	17	MG787419[b]

[a]*The ST of each isolate was determined by uploading the sequence of seven housekeeping genes to the S. epidermidis MLST online database (https://pubmlst.org/ sepidermidis/).* [b]*The genetic structure of these ACMEs has been described previously (McManus et al., 2017; O'Connor et al., 2018). ACME, arginine catabolic mobile element; ST, sequence type; OR, oral rinse; PP, periodontal pocket; SG, subgingival site.*

Statistical Analyses

In order to determine if the differences in the prevalence of staphylococcal species and isolates harboring ACME were significant between different sample sites or patient groups, two-tailed Fisher's exact tests were utilized. These analyses were carried out using GraphPad QuickCalcs[5]. A *P* value of <0.05 was deemed statistically significant. Statistical power analyses were calculated using the DSS research statistical power calculator tool[6] with a confidence interval of 5%.

Nucleotide Accession Numbers

The Genbank database accession numbers for the nucleotide sequences of the *S. epidermidis* ACMEs previously characterized (McManus et al., 2017; O'Connor et al., 2018) and in the present study are listed in **Table 1**.

[5]http://www.graphpad.com/quickcalcs/index.cfm

[6]https://www.dssresearch.com/KnowledgeCenter/toolkitcalculators/ statisticalpowercalculators.aspx

RESULTS

Prevalence of *S. epidermidis* and *S. aureus* in the Oral Cavity

Staphylococcus epidermidis was recovered from the oral rinse samples of 18/20 (90%) patients with periodontal disease, 18/38 (47.4%) patients with peri-implantitis, 25/31 (80.6%) orally healthy patients with implants and 44/64 (68.8%) orally healthy participants (**Table 2**). *Staphylococcus epidermidis* was significantly more prevalent in the oral rinse samples of orally healthy patients with implants than in those with peri-implantitis ($P = 0.0061$, Power = 90%), however the difference in the prevalence of *S. epidermidis* in the oral rinse samples of patients with periodontal disease in comparison to orally healthy participants was not quite statistically significant ($P = 0.081$).

The prevalence of *S. aureus* was considerably lower than *S. epidermidis* among all four groups of participants examined, detected in 5/20 (25%) patients with periodontal disease, 8/38 (21.1%) patients with peri-implantitis, 15/31 (48.4%) orally healthy patients with implants and 19/64 (29.7%) of orally healthy participants (**Table 2**). The prevalence of *S. aureus* was highest in the oral cavities of healthy patients with oral implants and was significantly more prevalent in the oral rinse samples of this patient group when compared to the corresponding sample sets from patients with peri-implantitis ($P = 0.0219$, Power = 77.8%).

Prevalence of *S. epidermidis* and *S. aureus* in Subgingival Sites, Peri-implant Sites and Periodontal Pockets

Staphylococcus epidermidis was recovered from the periodontal pockets of 6/20 (30%) patients with periodontal disease and the peri-implant sites of 15/29 (51.7%) patients with peri-implantitis. In contrast, *S. epidermidis* was only recovered from the subgingival sites of 5/31 (16.1%) orally healthy patients with implants and 5/64 (7.8%) of orally healthy participants (**Table 2**). *Staphylococcus epidermidis* was significantly more prevalent in the periodontal pockets of patients with periodontal disease than the subgingival sites of orally healthy participants ($P = 0.0189$, Power = 77.1%). Similarly, the prevalence of *S. epidermidis* was significantly higher in the subgingival sites of patients with peri-implantitis than similar sites in orally healthy patients with implants ($P = 0.0057$, Power = 91.4%).

Staphylococcus aureus was equally or less prevalent than *S. epidermidis* in the subgingival sites of all participant groups investigated, recovered from none of the periodontal pockets of patients with periodontal disease, 3/29 (10.3%) peri-implant pockets of patients with peri-implantitis, 4/31 (12.9%) of orally healthy patients with implants and 5/64 (7.8%) of orally healthy participants (**Table 2**). The prevalence of subgingival *S. aureus* was not significantly different in any of the four participant groups investigated.

TABLE 2 | Prevalence of ACME types harbored by *S. epidermidis* from distinct patient groups and anatomical sites.

Patients (n)	Sample site	Prevalence of *S. epidermidis* per patient (%)	Prevalence of *S. aureus* per patient (%)	Number of isolates harboring ACME (%)	Prevalence of *S. epidermidis* harboring ACME (%)	ACME types identified among *S. epidermidis* (n = patients)									
						I	II	III	IV	V	I and II	II and III	II and III	I and IV
Periodontal disease (20)[a]	OR	18/20 (90)	5/20 (25)	14/27 (51.9)	12/20 (60)	2	5	1	3	0	0	1	1	0
	PP	6/20 (30)[3]	0/20 (0)	6/9 (66.7)	5/20 (25)[7]	0	3	0	1	0	0	1	0	0
	SG	4/20 (20)	0/20 (0)	5/6 (83.3)	4/20 (20)	1	3	0	0	0	0	0	0	0
Peri-implantitis (29, 38)[b]	OR	18/38 (47.4)[1]	8/38 (21.1)[2]	4/15 (26.7)	4/38[c] (10.5)	0	3	1	0	0	0	0	0	0
	SG	15/29 (51.7)[4]	3/29 (10.3)	13/19 (68.4)	11/29 (37.9)[6,8]	0	8	1	1	1	0	3	0	0
Healthy with implants (31)	OR	25/31 (80.6)[1]	15/31 (48.4)[2]	24/39 (61.5)	19/31 (61.3)[5]	3	8	2	3	0	3	0	0	0
	SG	5/31 (16.1)	4/31 (12.9)	5/9 (55.6)	4/31 (12.9)[8]	1	0	1	1	1	0	0	0	1
Orally healthy (64)	OR	44/64 (68.8)	19/64 (29.7)	26/50 (52)	23/64 (35.9)[5]	2	15	1	3	0	0	0	2	0
	SG	5/64 (7.8)[3,4]	5/64 (7.8)	3/5 (60)	3/64 (4.7)[6,7]	0	1	0	2	0	0	0	0	0

Significant differences in the prevalence of S. epidermidis between paired sample types from distinct patient groups are indicated by identical numerals in superscript. [a]Subgingival samples were taken from patients with periodontal disease prior to any dental treatment. [b]Oral rinse samples were recovered from 38 patients with peri-implantitis; subgingival samples were taken from 29 of these patients. [c]Isolates recovered from oral rinse samples of four patients could not be located to be screened for the presence of ACME and thus were not included in statistical analyses. OR, oral rinse; SG, subgingival site; PP, periodontal pocket.

Prevalence of ACME Among *S. epidermidis* and *S. aureus* Isolates Recovered

The Arginine catabolic mobile element was detected in 100/179 (55.9%) of the *S. epidermidis* isolates recovered from all four participant groups (**Table 2**). The *mecA* gene was detected in 12/179 (6.7%) isolates; two from patients with periodontal disease, three from patients with peri-implantitis, three from orally healthy patients with implants and four from orally healthy participants. In total, 5/12 of the MRSE isolates identified also harbored ACMEs, predominantly type II.

Among the samples from which *S. epidermidis* was recovered, ACME was detected in isolates recovered from the oral rinse samples of 12/20 (60%) patients with periodontal disease, 4/38 (10.5%) patients with peri-implantitis, 19/31 (61.3%) orally healthy patients with implants, and 23/64 (35.9%) orally healthy participants. The prevalence of *S. epidermidis* isolates harboring ACME was significantly higher in the oral rinse samples of orally healthy patients with implants than in those of healthy participants ($P = 0.0275$, Power = 76.1%).

Although the prevalence of *S. epidermidis* was lower in subgingival sites than in the oral rinse samples of all four participant groups, the proportion of ACME-harboring *S. epidermidis* isolates was higher (**Table 2**). The presence of ACME was detected in *S. epidermidis* isolates from the periodontal pockets and subgingival sites of 5/20 (25%) and 4/20 (20%) patients with periodontal disease, respectively. Similarly, ACME was detected in 11/29 (37.9%), 4/31 (12.9%) and 3/64 (4.7%) *S. epidermidis* isolates from the subgingival sites of patients with peri-implantitis, healthy patients with dental implants and orally healthy participants, respectively.

Isolates harboring ACME were significantly more prevalent in subgingival samples of patients with peri-implantitis than in subgingival samples of orally healthy participants ($P = 0.0001$, Power = 98.4%). Similarly, isolates with ACME were also significantly more prevalent in periodontal pockets of patients with periodontal disease than subgingival sites of orally healthy participants ($P = 0.0167$, Power = 78.5%). Interestingly, the prevalence of ACME-harboring isolates was also significantly higher in the subgingival sites of patients with peri-implantitis than subgingival sites of orally healthy patients with implants ($P = 0.0369$, Power = 72.9%).

In contrast, ACME was not detected in any of the 83 *S. aureus* isolates recovered from oral rinse samples ($n = 56$) and subgingival samples ($n = 27$) investigated.

All five previously described ACME types were detected in *S. epidermidis* isolates in the present investigation by multiplex PCR. In all participant groups and anatomical sites sampled, ACME type II was the predominant ACME type, identified in *S. epidermidis* isolates recovered from 53 of the participants investigated. The recently described ACME types IV and V were detected in *S. epidermidis* isolates recovered from 15 and one participants sampled, respectively (**Table 2**).

Pairs of separate *S. epidermidis* isolates recovered from the oral rinse and subgingival samples of four orally healthy patients with implants were found to harbor distinct ACME types (i.e., types I and II or types I and IV). Similarly, a pair of *S. epidermidis* isolates recovered the same periodontal pocket of a patient with periodontal disease harbored ACME types I and II, respectively (**Table 2**). Furthermore, pairs of separate *S. epidermidis* isolates harboring ACMEs II and III were detected in the same oral rinse sample of one patient with periodontal disease and in two orally healthy participants.

A total of 35 *S. epidermidis* isolates selected as representatives of each ACME type, patient group, sample site and ACME type were subjected to WGS in order to elucidate the genetic organization of the ACMEs harbored in detail (**Table 1**).

Genetic Diversity of ACME Type I

Two isolates recovered from the oral rinse samples of a patient with periodontal disease (P14OR1) and an orally healthy patient with implants (I12OR1) were found to harbor ACME type I by multiplex PCR and were subjected to WGS. The STs of these isolates were identified as ST17 and ST7, respectively (**Table 1**). These STs are double locus variants of each other, differing at the *gtr* and *pyrR* loci by a total of four bp.

In both isolates, the ACME type I element was collocated with additional modules in CIs that were 54.3 and 39.9 kb in size, respectively (**Figure 1**). Both of the ACME type I elements characterized harbored the *arc* and *opp3* operons, the *speG* gene and were demarcated by DRs_B and _C (**Table 3**). In both isolates, ACME type I was located directly downstream of a module composed of the *copA* gene and the *ars* operon, separated by DR_B. In isolate I12OR1, this module was demarcated by an additional DR_B at the 5′ end in *orfX*, whereas in isolate P14OR1 this module was demarcated by DR_O at the 5′ end and an additional module containing three ORFs was located upstream of the *copA/ars* operon module (**Figure 1**).

Genetic Diversity of ACME Type II

Seventeen isolates harboring ACME type II were identified by multiplex PCR and were further investigated by WGS (**Figure 2**). All 17 isolates harbored ACMEs with the *arc* operon only and lacked both the *opp3* and *kdp* operons. Based on the differing combinations of DRs identified at the 5′ and 3′ ends, the ACME type II elements characterized could be divided into three distinct subtypes (IIa-c) of which ACME type IIb predominated (12/17, 70.6%). ACME type IIa was demarcated by the DRs _H and _C (**Table 3**) and was identified in 3/17 isolates (**Figures 2B–D**), all of which belonged to ST59. ACME type IIb (**Figures 2E–I**) was demarcated by the DRs _D and _C and was identified in 12/17 isolates belonging to ST89 ($n = 1$), ST14 ($n = 2$), ST73 ($n = 8$) and ST701, a single locus variant of ST73 ($n = 1$). The remaining two isolates harbored ACME type IIc demarcated by the DRs _A and _E and belonged to ST672 (**Figures 2J,K**). Interestingly in the case of the latter two isolates, the *copA* and *ars* operon were internalized within ACME type IIc in both isolates (**Figures 2J,K**) and both lacked the internal DR_G commonly identified downstream of the *arc* operon in ACME type II. The ACME type IIc-harboring CIs in these isolates were almost identical, differing only by the presence of an ORF upstream of the *sdrH* gene in isolate P14PPP2 (**Figure 2K**).
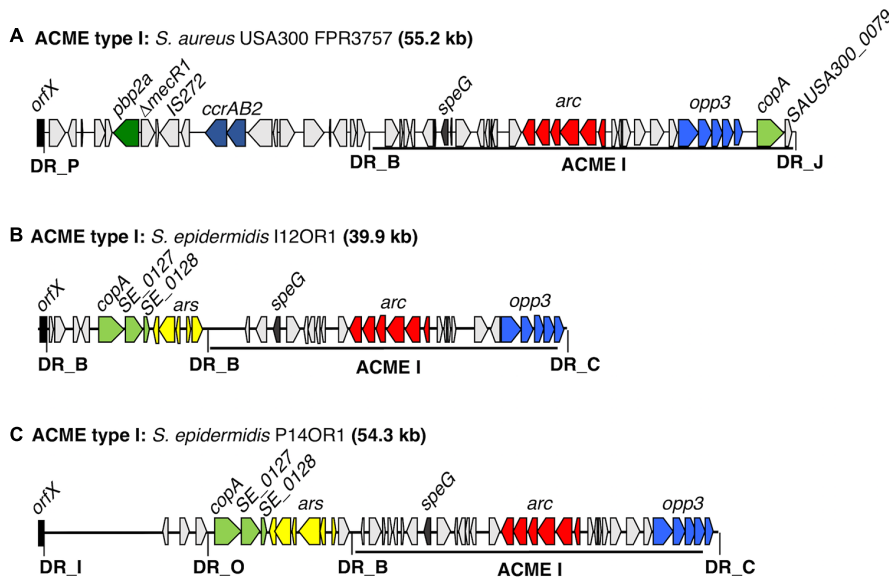
FIGURE 1 | Schematic representation of ACME type I elements harbored by two distinct *S. epidermidis* isolates investigated by WGS. The ACME type I previously described in the MRSA reference USA300 strain FPR3757 (Genbank accession number CP000255.1) is included for comparison **(A)**. The size of each ACME is indicated after each strain name. **(B,C)** Each gene or group of genes of interest is differentiated by a different color, i.e., light blue; *opp3*-operon, red; *arc*-operon, pale green; *cop*-operon, yellow; *ars* operon, dark blue; *ccrAB* complexes, dark green; *pbp2a*, dark gray; *speG*, light gray; genes encoding hypothetical proteins, sugar transporters, transposases and other ORFs, previously identified in ACMEs. The direction of transcription for each ORF is indicated by arrows. The DRs are indicated in bold font and correspond to DR sequences listed in **Table 3**.

TABLE 3 | Direct repeat sequences (DRs) identified among ACME types investigated.

DR	Sequence (5′-3′)	Associated ACME types/subtypes and CIs
DR_A	GAAGCATATCATAAATGA	IIa, IIb, IIc, IVa, V
DR_B	GAAGCGTATCACAAATAA	I, IIa, IIIª, IVa, V
DR_C	GAAGCGTATCGTAAGTGA	I, IIa, IIb, IIIª, IVa, IVb, V
DR_D	GAAGCGTACCACAAATAA	IIa, IIb,
DR_E	GAAGCGTATTAAAGTGAT	IIc
DR_F	GAAAGTTATCATAAGTGA	IVb, V
DR_G	GAAGCGTATAATAAGTAA	IIa, IIb, IIIª, IVa, IVb, V
DR_H	GAAGCGTATCATAAGTGA	IIa
DR_I	GAAGCGTATCATAAATGA	I
DR_J	GAGGCGTATCATAAGTAA	I
DR_L	GAAGCATATCATAAGTGA	IIa, IIIª, V
DR_M	GAAGGGTATCATAAATAA	IIIª
DR_N	GAAGCGTATCACAAATGA	IVa
DR_O	GAAGCATATCATAAATAA	I
DR_P	GAAGCTTATCATAAATGA	I

ªThe DRs _B, _L, _M, _G and _C described in the present study correspond to the DR1_A, DR_1B, DR_2, DR_3 and DR_4 previously described in ACME type III structures (McManus et al., 2017).

Five of the 17 ACME type II structures characterized by WGS existed as modules of larger CIs and were collocated with modules which harbored genes associated with SCC elements such as *ccr* and *mec* (**Figures 2B,C,I–K**). In each of these five CIs, the ACME type II structure was divided from these SCC-associated modules by DRs _A, _H, or _D (**Table 3**).

Interestingly, in one isolate (217PPP362) two separate and distinct SCC-associated modules were detected in tandem upstream of ACME type IIa (**Figure 2C**). The module immediately upstream of ACME type IIa was demarcated by the DRs _L and _H and harbored the *speG* and *ccrAB4* genes and was collocated immediately downstream of an additional SCC-associated module which harbored the *mecA* and *ccrAB2* genes. The ACME type IIa and IIb structures in the remaining 12/17 isolates investigated had integrated directly into *orfX* in the absence of any adjacent modules and were not components of larger CIs (**Figure 2**).

The presence of *sdr* genes, members of the serine/aspartate repeat family encoding fibrinogen-binding proteins was detected in modules adjacent to ACME type II in 3/17 isolates investigated and was collocated with ACME subtype IIb and IIc identified in the present study.

The *speG* gene encoding spermidine acetyltransferase was detected in only one isolate harboring ACME type II, (217PPP362) located in a SCC-associated module upstream of ACME type IIa (**Figure 2C**). The *copA* gene was located near the 3′ end of all 17 ACME type II structures investigated. In all ACME type IIa and IIb structures investigated, this gene was separated from the *arc* operon by DR_G and additional open reading frames commonly identified within ACMEs (**Figure 2**).

Genetic Diversity of ACME Type III

Four isolates harbored ACME type III, all of which were identified as ST329 (**Table 1**). Three of the ST329 isolates (P16OR1, 204OR1 and I11OR1) have been described previously (McManus
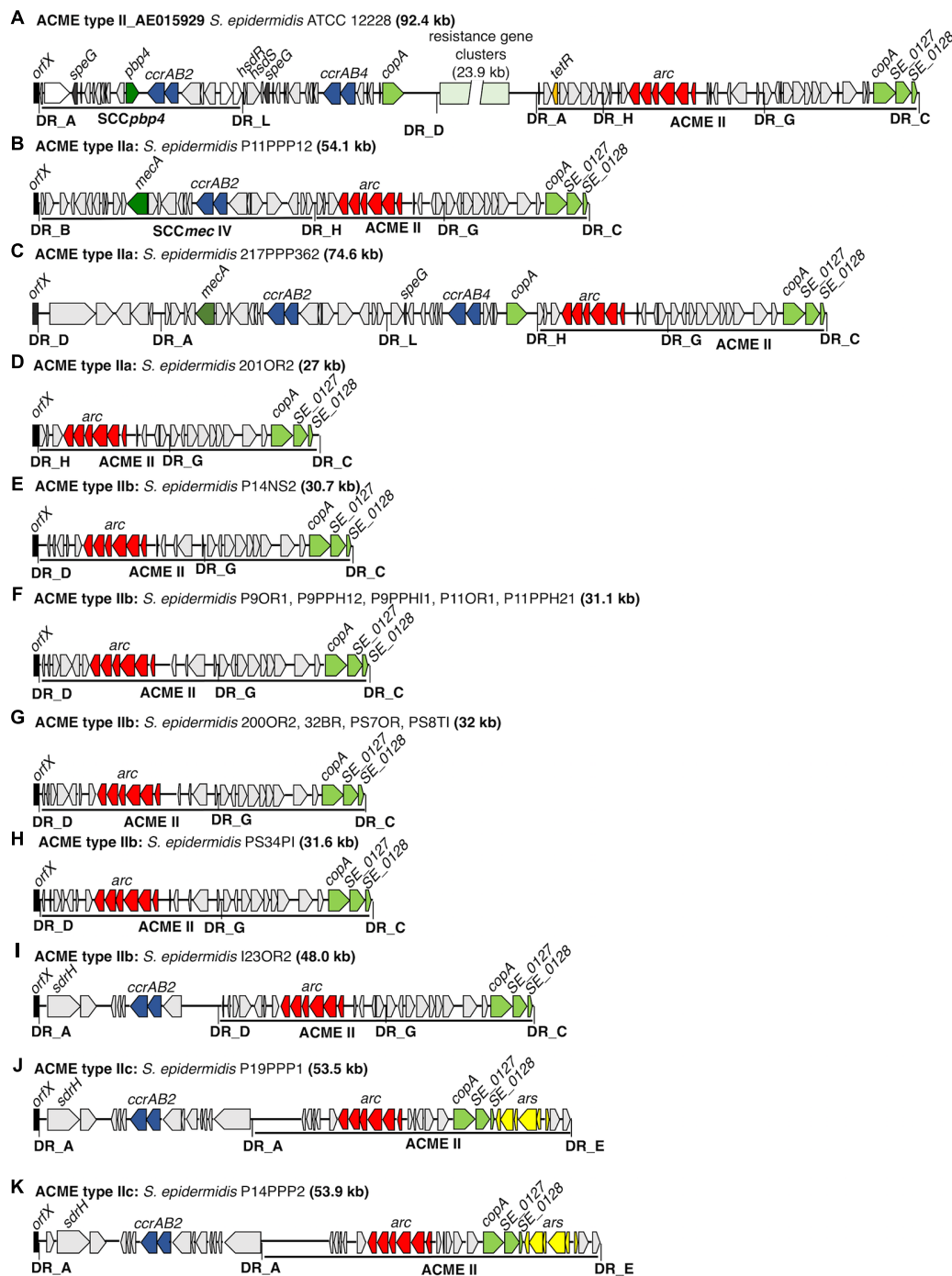
FIGURE 2 | Schematic representation of ACME II subtypes a-c and ACME type II-harboring CIs investigated in the present study. The previously described ACME type II CI harbored by the *S. epidermidis* reference strain ATCC12228 (Genbank accession number AE015929) is included for reference **(A)**. The size of each ACME is shown after the strain name. Each gene or group of genes of interest is differentiated by a different color, i.e., red; *arc*-operon, pale green; mustard; *tetR*, *cop*-operon, yellow; *ars* operon, dark blue; *ccrAB* genes, dark green; *mecA*, dark gray; *speG*, light gray; genes encoding hypothetical proteins, sugar transporters, transposases and other ORFs, previously identified in ACMEs. The direction of transcription for each ORF is indicated by the arrow. The DRs are indicated in bold font and correspond to each DR sequence listed in **Table 3**.

et al., 2017). The CI harbored by the fourth isolate (I1PPP121) consisted of two distinct modules separated by DR_G (**Figure 3**). The module at the 5′ end harbored two pairs of *ccr* genes, of which

ccrA4 and *ccrB2* were prematurely truncated. This module also harbored the *copA* and *ars* genes, located upstream of ACME type III as previously observed in the other three isolates. The ACME
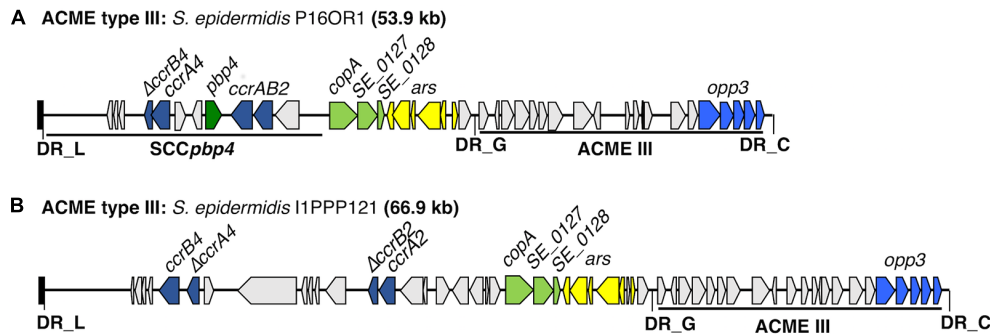
FIGURE 3 | Schematic representation of ACME type III structure from oral *S. epidermidis* isolate I1PPP121 characterized in the present study. The previously described ACME type III structure from an *S. epidermidis* isolate (P16OR1) recovered from an oral rinse sample from a patient with periodontal disease (McManus et al., 2017) is included for comparison **(A)**. The size of the ACME characterized is shown after the isolate name I1PPP121 **(B)**. Each gene or group of genes of interest is differentiated by a different color, i.e., pale green; *cop*-operon, dark blue; *ccrAB* gene complexes (of which the genes *ccrA4* and *ccrB2* were prematurely truncated), light blue; *opp3* operon, dark green; *pbp4*, light gray; genes encoding hypothetical proteins, sugar transporters, transposases and other ORFs, previously identified in ACMEs. The direction of transcription for each ORF is indicated by the arrow. The DRs are indicated in bold font and correspond to each DR sequence listed in **Table 3**.

type III harbored by I1PPP121 exhibited a minimum of 99.81% nucleotide identity to the ACME type IIIs harbored by isolates P16OR1, 204OR1, and I11OR1.

Genetic Diversity of ACME Types IV and V

Eleven isolates harboring ACME type IV were identified and belonged to ST210 ($n = 1$), ST153 ($n = 6$), ST130 ($n = 1$), ST17 ($n = 1$), ST297 ($n = 1$), and ST432 ($n = 1$) (**Table 1**). Nine of these isolates were described recently in a study that first defined ACME types IV and V in *S. epidermidis* (O'Connor et al., 2018). The two additional ST153 isolates (PS7P2 and PS23P1) were characterized in the present study and harbored ACME types IVa and IVb, respectively. The ACME type IVa harbored by isolate PS7P2 was identical to those previously described in isolates 120PPC, I9OR1 and I14OR1, and the ACME type IVb harbored by isolate PS23P1 was identical to that previously described in isolate P8OR3 (O'Connor et al., 2018). ACME type V was identified in only one isolate, recovered from the subgingival site of a patient with peri-implantitis and has been previously described (O'Connor et al., 2018).

Genetic Diversity Among the *arc*, *opp3* and *kdp* Operons Harbored by ACME Types I–V

The *arc* Operon

The percentage nucleotide identity between all ACME-*arc* operons identified in the present study, in ACME types I, II, IV and V ranged from 99.06 to 100%. The ACME-*arc* operons harbored by isolates 200OR2, 32BR, PS7OR and PS8TI (all ACME type IIb) exhibited 100% nucleotide identity to each other, as did that harbored by isolates 120PPC, I9OR1, I14OR1 (all ACME type IVa). Similarly, the ACME-*arc* operon harbored by P9OR1, P9PPH12, P9PPHI1, P11OR1 and P11PPH21 (all ACME type IIb) exhibited 100% nucleotide identity with each other (Supplementary Table 2).

The *opp3* Operon

The percentage nucleotide identity between all *opp3*-operons identified in the present study, in ACME types I, III and V ranged from 97.22 to 100% (Supplementary Table S3). The *opp3* operon harbored by isolates P16OR1 and P11OR1 (both ACME type III) exhibited 100% nucleotide identity to each other. The *opp3* operons harbored by ACME type I exhibited 99.91% to each other and 99.96% nucleotide homology to the *opp3* operon harbored by the reference ACME type I (Genbank accession number CP000255.1). The *opp3* operon harbored by the recently described ACME type V (O'Connor et al., 2018) exhibited 99.22-99.24% and 97.48–97.52% nucleotide identity to those harbored by ACME types III and I, respectively (Supplementary Table S3).

The *kdp* Operon

The *kdp* operon was highly conserved, exhibiting a minimum of 99.86% nucleotide identity amongst isolates harboring ACME types IV and V. The *kdp* operons harbored by isolates 120PPC, PS30PH, 33BR, 218PP361, PS7P2 and I9OR1 (all ACME type IVa) exhibited 100% nucleotide identity to each other, and to that harbored by PS23P1 (ACME type IVb). The *kdp* operon harbored by the remaining ACME type IVb isolates, PS36PD and P8OR3 exhibited 99.94% nucleotide identity to each other. The *kdp* operon harbored by ACME type V exhibited a minimum of 99.86% nucleotide identity to the *kdp* operon harbored by ACME types IVa and IVb (Supplementary Table S4).

Comparison of ACMEs Among Multiple *S. epidermidis* Isolates From the Same Patients

Three isolates recovered from the periodontal pockets and oral rinse sample of a patient with periodontal disease (P9) were all identified as ST73 (**Table 1**) and harbored genetically identical ACME type IIb structures (Supplementary Figure S1). In contrast, three distinct *S. epidermidis* isolates recovered from the oral rinse, nasal swab and periodontal pocket of a patient

(P14) with periodontal disease were identified as STs 17, 14 and 672, and harbored ACME types I, IIb and IIc, respectively (Supplementary Figure S1). Similarly, an isolate identified as ST59 and harboring ACME type IIa was recovered from the periodontal pocket of a patient with periodontal disease (P11) and was genetically distinct to two other isolates identified as ST73 and harboring ACME type IIb which were recovered from the oral rinse and another periodontal pocket of the same patient.

DISCUSSION

The Prevalence of Staphylococci in the Oral Cavity, Subgingival Sites and Periodontal Pockets

Previous investigations of the prevalence of staphylococcal species in the oral cavity are conflicting and/or ambiguous for several probable reasons. Firstly, studies did not definitively distinguish between distinct CoNS species and *S. aureus* (Leonhardt et al., 1999). Secondly, different studies used semi-discriminatory agar media such as Baird Parker or MSA for primary recovery, which may have resulted in failure to select and further distinguish between morphologically similar colonies of distinct CoNS species (Loberto et al., 2004; Cuesta et al., 2010; dos Santos et al., 2014). Thirdly, several previous studies relied on checkerboard DNA-DNA hybridization techniques for definitive identification of oral staphylococcal species from patients with dental implants, an approach that does not distinguish between viable and dead bacteria (Fürst et al., 2007; Renvert et al., 2008; Salvi et al., 2008). Subsequently, real-time PCR with species-specific primers demonstrated that the previously used DNA:DNA hybridization probes showed cross reactivity between *S. aureus* and *S. epidermidis* DNA (Cashin, 2013).

Patient sampling and recovery of viable isolates was undertaken using a uniform, systematic methodology for all sample sites and patient groups and identification of both *S. epidermidis* and *S. aureus* isolates was determined using robust procedures. The utilization of the chromogenic Sa*Select*^{TM} medium for primary isolation of oral staphylococci enabled direct visualization and presumptive identification of both *S. epidermidis* and *S. aureus* isolates based on the growth of these species as white/pale pink and pink colonies, respectively. This approach ensured that the differences in the prevalence of *S. epidermidis* and *S. aureus* observed are a true reflection of the patient groups and sample sites investigated.

Staphylococcus epidermidis
Staphylococcus epidermidis was detected in the oral cavities of 47.4 – 90% of the four participant groups investigated (**Table 2**) which was higher than previous studies that reported its recovery from the oral cavities of 27.3% patients with periodontal disease (Loberto et al., 2004) and 41% of orally healthy participants (Ohara-Nemoto et al., 2008).

Previous reports of the prevalence of subgingival *S. epidermidis* vary greatly, ranging between 15.9 and 64.3% in periodontal pockets (Loberto et al., 2004; Murdoch et al.,

2004; dos Santos et al., 2014) and between 42.9 and 60.7% in the subgingival sites of healthy participants (Murdoch et al., 2004; Ohara-Nemoto et al., 2008), most likely due to different methods used. Data on the subgingival prevalence of *S. epidermidis* in patients with oral implants and/or peri-implantitis are largely lacking or did not undertake definitive identification of this species (Leonhardt et al., 1999).

Staphylococcus aureus
In the present study, *S. aureus* was considerably less prevalent than *S. epidermidis* in the oral cavities of participant groups and was rarely detected in subgingival sites or periodontal pockets (**Table 2**). The dramatic contrast in the oral prevalence of *S. epidermidis* and *S. aureus* was not surprising, as a negative correlation between the prevalence of these species has been reported previously (Brescó et al., 2017). Previous reports of the prevalence of subgingival *S. aureus* vary greatly, ranging from 13.4 to 68.2% (Cuesta et al., 2010; Zhuang et al., 2014) in periodontal pockets, 40–70.4% of peri-implant pockets and 25–44% of healthy subgingival sites around dental implants (Fürst et al., 2007; Renvert et al., 2008; Salvi et al., 2008). The contrasting results between the present and previous studies most likely reflects differences in methodology used as discussed above.

The Prevalence of ACME in *S. epidermidis* and *S. aureus*
The prevalence of ACME has previously been investigated amongst *S. epidermidis* populations and has been reported to range from 40 to 65.4% in MRSE, and from 64.4 to 83% in methicillin susceptible *S. epidermidis* (Miragaia et al., 2009; Barbier et al., 2011; Du et al., 2013; Onishi et al., 2013). In the present investigation, only 12/179 (6.7%) of the *S. epidermidis* isolates recovered were MRSE. This finding correlates with previous studies that reported a higher prevalence of ACME in methicillin-susceptible *S. epidermidis*. Five of the 12 MRSE isolates identified here harbored ACMEs, predominantly ACME type II.

This is the first investigation into the prevalence of ACME amongst *S. epidermidis* and *S. aureus* isolates recovered from both above and below the gumline in patients with and without oral disease. The prevalence of ACME was greater than 50% amongst the populations of *S. epidermidis* recovered. In contrast, ACME was not detected in any of the *S. aureus* isolates recovered.

This study revealed for the first time that the prevalence of ACME-harboring *S. epidermidis* was significantly higher in subgingival sites of patients with peri-implantitis and periodontal disease in comparison to healthy individuals ($P = 0.0001$ and $P = 0.0167$, respectively). Furthermore, *S. epidermidis* harboring ACME were significantly more prevalent in the subgingival sites of patients with peri-implantitis than in orally healthy patients with implants. Together, these results suggest a strong association of *S. epidermidis* isolates harboring ACME with diseased, semi-anaerobic subgingival tissue sites. No correlation between specific ACME types or subtypes and specific disease state or oral site was identified.

One of the potential limitations of the present investigation was the number of patients sampled. The study group was limited to patients attending DDUH who did not have any underlying conditions and had not received antibiotics in the 2 months prior to sampling. It is likely that larger investigations would further support the findings of this study as well as likely identify additional ACME types. Furthermore, it would be interesting to determine if ACME is more abundant in *S. epidermidis* isolates recovered from other diseased anatomical sites or wounds.

The Genetic Diversity of ACME

Previous investigations of the prevalence and structural diversity of ACME relied on primers targeting the ACME-*arc* and ACME-*opp3* genes to detect ACME types I–III only (Miragaia et al., 2009; Onishi et al., 2013). This study is the first to include primers targeting the recently described ACME-*kdp* operon in a multiplex ACME-typing PCR and therefore the results of this study accurately reflect the true prevalence of ACME types currently described in *S. epidermidis*, at least of oral origin. The application of WGS to the characterization of ACME has revealed the additional ACME types harbored by *S. epidermidis* in addition to the detailed structural diversity of these ACMEs. It is highly likely that *S. epidermidis* isolates harboring the *opp3* and *kdp* operons and the *kdp* operon only will also be identified in the future using WGS.

Common DRs were often observed amongst multiple ACME types and subtypes (**Table 3**). Many distinct DRs have been identified at the 5′ end of ACME, which is clearly demarcated by the integration of this element into *orfX*. The demarcation of the 3′ end is less obvious. Previously described ACMEs terminated at the 3′ end by DR_J in ACME type I of *S. aureus* (Genbank accession number CP000255.1) (Diep et al., 2006) and by DR_C in ACME types II–V of *S. epidermidis* (Gill et al., 2005; McManus et al., 2017; O'Connor et al., 2018), even though an internal DR_G is present downstream of the *arc* operon in both the reference ACME type II (Genbank accession number AE015929) and subsequently described in ACME types II (**Figure 2**) and IV (O'Connor et al., 2018). Many of the ACMEs characterized in the present and previous studies were components of CIs that were separated by multiple DRs, and greatly contributed to the diversity of these MGEs. Interestingly, the *copA* gene and *ars* operon were commonly detected in various distinct positions within or adjacent to ACME types I, II, III, and V. These genes were identified downstream of the ACME type II element in several isolates investigated in the present study (data not shown) but were internalized within the ACME type IIc structures (**Figure 2**). The identification of multiple DRs in common amongst distinct ACME types and the frequent organizational differences observed both within and between each ACME type supports previous studies that suggested ACMEs are assembled in a stepwise, modular fashion (Thurlow et al., 2013). Indeed, ACME diversity was also observed among separate isolates recovered from the oral cavities of the same patient (Supplementary Figure S1) in three cases.

The Association of ACME With Specific *S. epidermidis* Lineages

Isolates harboring ACME type I investigated here were identified as STs 7 and 17 (**Table 1**), both of which have been previously assigned to GC6 by Bayesian clustering analysis (Thomas et al., 2014; Tolo et al., 2016), a GC that is enriched with ACME-harboring isolates and associated with infection sources. Isolates harboring ACME type II were identified as STs 73 (and ST701, a single locus variant of ST73), 59, 89, 14, and 672, four of which have previously been assigned to GC1, a GC enriched with isolates from non-hospital sources (Thomas et al., 2014; Tolo et al., 2016). Two isolates harboring ACME type IIc from periodontal pockets of two separate patients were identified as ST672, however, this ST was not previously assigned to a GC. All isolates harboring ACME type III were recovered from separate patients and all belonged to ST329, previously associated with GC4. Isolates belonging to GC4 have been associated with a more commensal lifestyle (Thomas et al., 2014). Isolates harboring ACME type IV were identified as STs 153, 297, 130, and 17, of which four were previously assigned to GC6. The isolate harboring ACME type V, identified as ST5, also belonged to GC6.

Distinct ACME types were more commonly associated with isolates belonging to identical or closely related STs, rather than the participant group or sample sites from which each ACME-harboring isolate was recovered. Indeed, many STs identified amongst isolates recovered from the oral rinse samples of healthy participants belonged to GC6 (**Table 2**), a cluster enriched with isolates from infections (Tolo et al., 2016). These findings strongly suggest that the stepwise accumulation of ACMEs occurs in specific lineages of *S. epidermidis*, rather than in specific anatomical sites.

The Potential Role of ACME in Disease

In the present study, the predominant ACME types detected (II and IV) in *S. epidermidis* harbored the *arc* operon. Researchers have hypothesized that ACME enhances the transmissibility, colonization and persistence of the MRSA USA300 strain on human skin, contributing to the success of this lineage (Diep et al., 2008; Planet et al., 2013). The arginine deaminase pathway encoded by *arc* is responsible for the breakdown of polyamines which are largely L-arginine based. This results in the formation of ornithine, ATP, CO_2 and ammonia, the latter of which contributes to the internal pH regulation of staphylococci in acidic conditions (Planet et al., 2013; Lindgren et al., 2014). The contribution of the constitutively expressed *arc* operon is likely to be highly advantageous to staphylococcal survival in the acidic environments present in dental plaque. In addition, the ATP generated is likely beneficial to organisms living in nutrient poor, semi-anaerobic environments such as present in periodontal pockets. Polyamines are associated with biological processes such as wound healing and infection clearance and it is therefore likely that they would be highly associated with oral inflammatory diseases such as periodontal disease and peri-implantitis. The *speG* gene, encoding a spermidine acetyltransferase, is thought

to mitigate the lethal effects of polyamines on staphylococci. The exact benefit of the *opp3* operon is unclear but encodes an oligopeptide permease ABC transporter (Diep et al., 2006, 2008). The *speG* gene and *opp3* operon were detected in only eight and seven of the 35 ACMEs characterized, respectively, suggesting that these genes are relatively dispensable for *S. epidermidis* in oral environments (McManus et al., 2017; O'Connor et al., 2018).

Interestingly, the *kdp* operon was detected in 12 of the 35 ACMEs characterized by WGS, suggesting that these genes contribute to the survival of *S. epidermidis* in oral environments. This operon encodes a potassium transporter that is important for maintaining intracellular pH homeostasis and metabolic processes in *S. aureus* (Price-Whelan et al., 2013) and likely plays an important role in the adaptation and survival of *S. epidermidis* in dental plaque, which has a significant concentration of K^+ ions (Margolis and Moreno, 1994).

CONCLUSION

This study revealed the significantly high prevalence of *S. epidermidis* in periodontal pockets and subgingival sites of patients with periodontal disease or peri-implantitis, respectively. There was also a very significant difference in the prevalence of *S. epidermidis* harboring ACME in these diseased subgingival sites and periodontal pockets compared to those recovered from healthy subgingival sites and oral rinse samples (**Table 2**). As yet, it is unclear if this organism contributes to disease progression directly. The presence of five main ACME types among oral *S. epidermidis* isolates was identified and extensive genetic diversity among these types was revealed using WGS, which would have been overlooked using previously described multiplex PCRs (Miragaia et al., 2009; Onishi et al., 2013). The *arc* and *kdp* operons harbored by the predominant ACME types identified (II and IV) very likely contribute to the survival of oral *S. epidermidis* under diseased and inflammatory conditions such as periodontal disease and peri-implantitis.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the St. James's Hospital and Federated Dublin Voluntary Hospitals Joint Research Ethics Committee (JREC) and the Faculty of Health Sciences Ethics Committee of Trinity College Dublin, Ireland. The protocol and all documentation (including consent forms) provided to patients was pre-approved by the Research Ethics Committees. Prior to enrollment in the study, all participants were provided with comprehensive patient information documentation and provided written consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

AO' conceived and designed the study, performed the WGS data analysis, and drafted the manuscript. BM conceived and designed the study and helped with the study co-ordination, WGS data analysis, and wrote the manuscript. PK assisted with bioinformatic analyses. GB and TF performed definitive identification of staphylococcal isolates by MALDI_TOF. PC assisted with laboratory processing and isolation of staphylococcal isolates. MOS and IP performed clinical sampling of patients and participants included in the study. DC conceived and designed the study, purchased the required materials, assisted with data analysis, and drafted the manuscript. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01558/full#supplementary-material

REFERENCES

Adell, R. (1981). A 15-year study of osseointegrated implants in the treatment of the edentulous jaw. *Int. J. Oral Surg.* 10, 387–416. doi: 10.1016/S0300-9785(81)80077-4

Baig, S., Johannesen, T. B., Overballe-Petersen, S., Larsen, J., Larsen, A. R., and Stegger, M. (2018). Novel SCC*mec* type XIII (9A) identified in an ST152 methicillin-resistant *Staphylococcus aureus*. *Infect. Genet. Evol.* 61, 74–76. doi: 10.1016/j.meegid.2018.03.013

Barbier, F., Lebeaux, D., Hernandez, D., Delannoy, A. S., Caro, V., François, P., et al. (2011). High prevalence of the arginine catabolic mobile element in carriage isolates of methicillin-resistant *Staphylococcus epidermidis*. *J. Antimicrob. Chemother.* 66, 29–36. doi: 10.1093/jac/dkq410

Berriman, M., and Rutherford, K. (2003). Viewing and annotating sequence data with Artemis. *Br. Bioinform.* 4, 124–132.

Branemark, P. I., Adell, R., Albrektsson, T., Lekholm, U., Lundkvist, S., and Rockler, B. (1983). Osseointegrated titanium fixtures in the treatment of edentulousness. *Biomaterials* 4, 25–28.

Brescó, M. S., Harris, L. G., Thompson, K., Stanic, B., Morgenstern, M., O'Mahony, L., et al. (2017). Pathogenic mechanisms and host interactions in *Staphylococcus epidermidis* device-related infection. *Front. Microbiol.* 8:1401. doi: 10.3389/fmicb.2017.01401

Cashin, P. J. (2013). *Molecular Characterisation of Staphylococcus aureus and Staphylococcus epidermidis Populations from Healthy and Diseased Oral Implant and Natural Tooth Sites in a Cohort of Patients with Periimplantitis.* Available at: http://hdl.handle.net/2262/77912

Cuesta, A. I., Jewtuchowicz, V., Brusca, M. I., Nastri, M. L., and Rosa, A. C. (2010). Prevalence of *Staphylococcus* spp and *Candida* spp in the oral cavity and periodontal pockets of periodontal disease patients. *Acta Odontol. Latinoam.* 23, 20–26.

Diep, B. A., Gill, S. R., Chang, R. F., Phan, T. H., Van Chen, J. H., Davidson, M. G., et al. (2006). Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus. Lancet* 367, 731–739. doi: 10.1016/S0140-6736(06)68 231-7

Diep, B. A., Stone, G. G., Basuino, L., Graber, C. J., Miller, A., des Etages, S. A., et al. (2008). The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus. J. Infect. Dis.* 197, 1523–1530. doi: 10.1086/587907

dos Santos, B. R., Demeda, C. F., da Silva, E. E., de Britto, M. H., Lima, K. C., and de Melo, M. C. (2014). Prevalence of subgingival *Staphylococcus* at periodontally healthy and diseased sites. *Braz. Dent. J.* 25, 271–276. doi: 10.1590/0103-6440201302285

Du, X., Zhu, Y., Song, Y., Li, T., Luo, T., Sun, G., et al. (2013). Molecular analysis of *Staphylococcus epidermidis* strains isolated from community and hospital environments in China. *PLoS One* 8:e62742. doi: 10.1371/journal.pone.0062742

Espedido, B. A., Steen, J. A., Barbagiannakos, T., Mercer, J., Paterson, D. L., Grimmond, S. M., et al. (2012). Carriage of an ACME II variant may have contributed to methicillin- resistant *Staphylococcus aureus* sequence type 239-like strain replacement in Liverpool Hospital, Sydney, Australia. *Antimicrob. Agents Chemother.* 56, 3380–3383. doi: 10.1128/AAC.00013-12

Friedlander, A. H. (2010). Oral cavity staphylococci are a potential source of prosthetic joint infection. *Clin. Infect. Dis.* 50, 1682–3; author reply 1683. doi: 10.1086/653003

Fürst, M. M., Salvi, G. E., Lang, N. P., and Persson, G. R. (2007). Bacterial colonization immediately after installation on oral titanium implants. *Clin. Oral Implants Res.* 18, 501–508. doi: 10.1111/j.1600-0501.2007.01381.x

Gill, S. R., Fouts, D. E., Archer, G. L., Mongodin, E. F., Deboy, R. T., Ravel, J., et al. (2005). Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J. Bacteriol.* 187, 2426–2438. doi: 10.1128/JB.187.7.2426-2438.2005

Hajishengallis, G. (2015). Periodontitis: from microbial immune subversion to systemic inflammation. *Nat. Rev. Immunol.* 15, 30–44. doi: 10.1038/nri3785

Holtfreter, B., Albandar, J. M., Dietrich, T., Dye, B. A., Eaton, K. A., Eke, P. I., et al. (2015). Standards for reporting chronic periodontitis prevalence and severity in epidemiologic studies: proposed standards from the Joint EU/USA Periodontal Epidemiology Working Group. *J. Clin. Periodontol.* 42, 407–412. doi: 10.1111/jcpe.12392

Hon, P. Y., Koh, T. H., Tan, T. Y., Krishnan, P., Leong, J. W. Y., Jureen, R., et al. (2014). Changing molecular epidemiology and high rates of mupirocin resistance among meticillin-resistant *Staphylococcus aureus* in Singaporean hospitals. *J. Glob. Antimicrob. Resist.* 2, 53–55. doi: 10.1016/j.jgar.2013.10.002

International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (2009). Classification of staphylococcal cassette chromosome *mec* (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrob. Agents Chemother.* 53, 4961–4967. doi: 10.1128/AAC.00579-09

Ito, T., Katayama, Y., Asada, K., Mori, N., Tsutsumimoto, K., Tiensasitorn, C., et al. (2001). Structural comparison of three types of staphylococcal cassette chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus. Antimicrob. Agents Chemother.* 45, 1323–1336. doi: 10.1128/AAC.45.5.1323-1336.2001

Kawaguchiya, M., Urushibara, N., Ghosh, S., Kuwahara, O., Morimoto, S., Ito, M., et al. (2013). Genetic diversity of emerging Panton-Valentine leukocidine/arginine catabolic mobile element (ACME)-positive ST8 SCCmec-IVa meticillin-resistant *Staphylococcus aureus* (MRSA) strains and ACME-positive CC5 (ST5/ST764) MRSA strains in northern Japan. *J. Med. Microbiol.* 62, 1852–1863. doi: 10.1099/jmm.0.062125-0

Leonhardt, Å., Renvert, S., and Dahlén, G. (1999). Microbial findings at failing implants. *Clin. Oral Implants Res.* 10, 339–345. doi: 10.1034/j.1600-0501.1999.100501.x

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Lindgren, J. K., Thomas, V. C., Olson, M. E., Chaudhari, S. S., Nuxoll, A. S., Schaeffer, C. R., et al. (2014). Arginine deiminase in *Staphylococcus epidermidis* functions to augment biofilm maturation through pH homeostasis. *J. Bacteriol.* 196, 2277–2289. doi: 10.1128/jb.00051-14

Loberto, J. C. S., De Paiva Martins, C. A., Ferreira Dos Santos, S. S., Cortelli, J. R., and Cardoso Jorge, A. O. (2004). *Staphylococcus* spp. in the oral cavity and periodontal pockets of chronic periodontitis patients. *Braz. J. Microbiol.* 35, 64–68. doi: 10.1590/S1517-83822004000100010

Margolis, H. C., and Moreno, E. C. (1994). Composition and cariogenic potential of dental plaque fluid. *Crit. Rev. Oral Biol. Med.* 5, 1–25.

McManus, B. A., O'Connor, A. M., Kinnevey, P. M., O'Sullivan, M., Polyzois, I., and Coleman, D. C. (2017). First detailed genetic characterization of the structural organization of type III arginine catabolic mobile elements harbored by *Staphylococcus epidermidis* by using whole-genome sequencing. *Antimicrob. Agents Chemother.* 61, e1216-17. doi: 10.1128/AAC.01216-17

Miragaia, M., de Lencastre, H., Perdreau-Remington, F., Chambers, H. F., Higashi, J., Sullam, P. M., et al. (2009). Genetic diversity of arginine catabolic mobile element in *Staphylococcus epidermidis. PLoS One* 4:e7722. doi: 10.1371/journal.pone.0007722

Monecke, S., Jatzwauk, L., Weber, S., Slickers, P., and Ehricht, R. (2008). DNA microarray-based genotyping of methicillin-resistant *Staphylococcus aureus* strains from Eastern Saxony. *Clin. Microbiol. Infect.* 14, 534–545. doi: 10.1111/j.1469-0691.2008.01986.x

Murdoch, F. E., Sammons, R. L., and Chapple, I. L. C. (2004). Isolation and characterization of subgingival staphylococci from periodontitis patients and controls. *Oral Dis.* 10, 155–162. doi: 10.1046/j.1601-0825.2003.01000.x

Nibali, L., Pometti, D., Chen, T. T., and Tu, Y. K. (2015). Minimally invasive non-surgical approach for the treatment of periodontal intrabony defects: a retrospective analysis. *J. Clin. Periodontol.* 42, 853–859. doi: 10.1111/jcpe.12443

O'Connor, A. M., McManus, B. A., and Coleman, D. C. (2018). First description of novel arginine catabolic mobile elements (ACMEs) types IV and V harboring a *kdp* operon in *Staphylococcus epidermidis* characterized by whole genome sequencing. *Infect. Genet. Evol.* 61, 60–66. doi: 10.1016/j.meegid.2018.03.012

Ohara-Nemoto, Y., Haraga, H., Kimura, S., and Nemoto, T. K. (2008). Occurrence of staphylococci in the oral cavities of healthy adults and nasal-oral trafficking of the bacteria. *J. Med. Microbiol.* 57, 95–99. doi: 10.1099/jmm.0.47561-0

Onishi, M., Urushibara, N., Kawaguchiya, M., Ghosh, S., Shinagawa, M., Watanabe, N., et al. (2013). Prevalence and genetic diversity of arginine catabolic mobile element (ACME) in clinical isolates of coagulase-negative staphylococci: identification of ACME type I variants in *Staphylococcus epidermidis. Infect. Genet. Evol.* 20, 381–388. doi: 10.1016/j.meegid.2013.09.018

Otto, M. (2013). Coagulase-negative staphylococci as reservoirs of genes facilitating MRSA infection: Staphylococcal commensal species such as *Staphylococcus epidermidis* are being recognized as important sources of genes promoting MRSA colonization and virulence. *Bioessays* 35, 4–11. doi: 10.1002/bies.201200112

Percival, S. L., McCarty, S., Hunt, J. A., and Woods, E. J. (2014). The effects of pH on wound healing, biofilms, and antimicrobial efficacy. *Wound Repair Regen.* 22, 174–186. doi: 10.1111/wrr.12125

Planet, P. J., LaRussa, S. J., Dana, A., Smith, H., Xu, A., Ryan, C., et al. (2013). Emergence of the epidemic methicillin-resistant *Staphylococcus aureus* strain USA300 coincides with horizontal transfer of the arginine catabolic mobile element and *speG*-mediated adaptations for survival on skin. *mBio* 4:e889-13. doi: 10.1128/mBio.00889-13

Price-Whelan, A., Poon, C. K., Benson, M. A., Eidem, T. T., Roux, C. M., Boyd, J. M., et al. (2013). Transcriptional profiling of *Staphylococcus aureus* during growth in 2 M NaCl leads to clarification of physiological roles for Kdp and Ktr K+ uptake systems. *mBio* 4:e407-13. doi: 10.1128/mBio.00407-13

Renvert, S., Lindahl, C., Renvert, H., and Persson, G. R. (2008). Clinical and microbiological analysis of subjects treated with Branemark or Astra Tech implants: a 7-year follow-up study. *Clin. Oral Implants Res.* 19, 342–347. doi: 10.1111/j.1600-0501.2007.01476.x

Renvert, S., and Polyzois, I. (2018). Treatment of pathologic peri-implant pockets. *Periodontol* 2000, 180–190. doi: 10.1111/prd.12149

Rolo, J., de Lencastre, H., and Miragaia, M. (2012). Strategies of adaptation of *Staphylococcus epidermidis* to hospital and community: amplification and diversification of SCC*mec. J. Antimicrob. Chemother.* 67, 1333–1341. doi: 10.1093/jac/dks068

Salvi, G. E., Fürst, M. M., Lang, N. P., and Persson, G. R. (2008). One-year bacterial colonization patterns of *Staphylococcus aureus* and other bacteria at implants and adjacent teeth. *Clin. Oral Implants Res.* 19, 242–248. doi: 10.1111/j.1600-0501.2007.01470.x

Sanz, M., and Chapple, I. L. (2012). Clinical research on peri-implant diseases: consensus report of Working Group 4. *J. Clin. Periodontol.* 39, 202–206. doi: 10.1111/j.1600-051X.2011.01837.x

Shore, A. C., and Coleman, D. C. (2013). Staphylococcal cassette chromosome *mec*: recent advances and new insights. *Int. J. Med. Microbiol.* 303, 350–359. doi: 10.1016/j.ijmm.2013.02.002

Shore, A. C., Rossney, A. S., Brennan, O. M., Kinnevey, P. M., Humphreys, H., Sullivan, D. J., et al. (2011). Characterization of a novel arginine catabolic mobile element (ACME) and staphylococcal chromosomal cassette *mec* composite island with significant homology to *Staphylococcus epidermidis* ACME type II in methicillin-resistant *Staphylococcus aureus* genotype. *Antimicrob. Agents Chemother.* 55, 1896–1905. doi: 10.1128/AAC.01756-10

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Song, Z., Borgwardt, L., Hoiby, N., Wu, H., Sorensen, T. S., and Borgwardt, A. (2013). Prosthesis infections after orthopedic joint replacement: the possible role of bacterial biofilms. *Orthop. Rev.* 5, 65–71. doi: 10.4081/or.2013.e14

Soroush, S., Jabalameli, F., Taherikalani, M., Amirmozafari, N., Imani Fooladi, A. A., Asadollahi, K., et al. (2016). Investigation of biofilm formation ability, antimicrobial resistance and the staphylococcal cassette chromosome *mec* patterns of methicillin resistant *Staphylococcus epidermidis* with different sequence types isolated from children. *Microb. Pathog.* 93, 126–130. doi: 10.1016/j.micpath.2016.01.018

Thomas, J. C., Vargas, M. R., Miragaia, M., Peacock, S. J., Archer, G. L., and Enright, M. C. (2007). Improved multilocus sequence typing scheme for *Staphylococcus epidermidis. J. Clin. Microbiol.* 45, 616–619. doi: 10.1128/jcm.01934-06

Thomas, J. C., Zhang, L., and Robinson, D. A. (2014). Differing lifestyles of *Staphylococcus epidermidis* as revealed through Bayesian clustering of multilocus sequence types. *Infect. Genet. Evol.* 22, 257–264. doi: 10.1016/j.meegid.2013.06.020

Thurlow, L. R., Joshi, G. S., Clark, J. R., Spontak, J. S., Neely, C. J., Maile, R., et al. (2013). Functional modularity of the arginine catabolic mobile element contributes to the success of USA300 methicillin-resistant *Staphylococcus aureus. Cell Host Microbe* 13, 100–107. doi: 10.1016/j.chom.2012.11.012

Thurnheer, T., and Belibasakis, G. N. (2016). Incorporation of staphylococci into titanium-grown biofilms: an *in vitro* "submucosal" biofilm model for peri-implantitis. *Clin. Oral Implants Res.* 27, 890–895. doi: 10.1111/clr.12715

Tolo, I., Thomas, J. C., Fischer, R. S. B., Brown, E. L., Gray, B. M., and Robinson, D. A. (2016). Do *Staphylococcus epidermidis* genetic clusters predict isolation sources? *J. Clin. Microbiol.* 54, 1711–1719. doi: 10.1128/JCM.03345-15

Wu, Z., Li, F., Liu, D., Xue, H., and Zhao, X. (2015). Novel type XII staphylococcal cassette chromosome *mec* harboring a new cassette chromosome recombinase, CcrC2. *Antimicrob. Agents Chemother.* 59, 7597–7601. doi: 10.1128/AAC.01692-15

Xue, T., You, Y., Hong, D., Sun, H., and Sun, B. (2011). The *Staphylococcus aureus* KdpDE two-component system couples extracellular K+ sensing and Agr signaling to infection programming. *Infect. Immun.* 79, 2154–2167. doi: 10.1128/IAI.01180-10

Zhuang, L. F., Watt, R. M., Mattheos, N., Si, M. S., Lai, H. C., and Lang, N. P. (2014). Periodontal and peri-implant microbiota in patients with healthy and inflamed periodontal and peri-implant tissues. *Clin. Oral Implants Res.* 27, 13–21. doi: 10.1111/clr.12508

Factors Contributing to the Evolution of *mecA*-Mediated β-lactam Resistance in Staphylococci: Update and New Insights From Whole Genome Sequencing (WGS)

Maria Miragaia*

Laboratory of Bacterial Evolution and Molecular Epidemiology, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

The understanding of the mechanisms of antibiotic resistance development are fundamental to alert and preview beforehand, the large scale dissemination of resistance to antibiotics, enabling the design of strategies to prevent its spread. The *mecA*-mediated methicillin resistance conferring resistance to broad-spectrum β-lactams is globally spread in staphylococci including hospitals, farms and community environments, turning ineffective the most widely used and efficient class of antibiotics to treat staphylococcal infections. The use of whole genome sequencing (WGS) technologies at a bacterial population level has provided a considerable progress in the identification of key steps that led to *mecA*-mediated β-lactam resistance development and dissemination. Data obtained from multiple studies indicated that *mecA* developed from a harmless core gene (*mecA1*) encoding the penicillin-binding protein D (PbpD) from staphylococcal species of animal origin (*S. sciuri* group) due to extensive β-lactams use in human created environments. Emergence of the resistance determinant involved distortion of PbpD active site, increase in *mecA1* expression, addition of regulators (*mecR1*, *mecI*) and integration into a mobile genetic element (SCC*mec*). SCC*mec* was then transferred into species of coagulase-negative staphylococci (CoNS) that are able to colonize both animals and humans and subsequently transferred to *S. aureus* of human origin. Adaptation of *S. aureus* to the exogenously acquired SCC*mec* involved, deletion and mutation of genes implicated in general metabolism (auxiliary genes) and general stress response and the adjustment of metabolic networks, what was accompanied by an increase in β-lactams minimal inhibitory concentration and the transition from a heterogeneous to homogeneous resistance profile. Nowadays, methicillin-resistant *S. aureus* (MRSA) carrying SCC*mec* constitutes one of the most important worldwide pandemics. The stages of development of *mecA*-mediated β-lactam resistance described here may serve as a model for previewing and preventing the emergence of resistance to other classes of antibiotics.

Keywords: β-lactams resistance, *Staphylococcus sciuri*, staphylococcal cassette chromosome *mec* (SCC*mec*), methicillin-resistant *Staphylococcus aureus* (MRSA), whole genome sequencing

INTRODUCTION

Antimicrobial resistance threatens the effective prevention and treatment strategies of an increasing range of bacterial infections. In 21st century we are facing the real possibility that minor injuries and common infections can lead to death. A detailed understanding of the evolutionary processes occurring in nature that lead to resistance development is thus essential for anticipating its emergence and to restrain its spread.

One of the best models of resistance development is the emergence of methicillin resistance in staphylococci not only due the fact that it is extremely well documented, but mainly because it gave rise to methicillin-resistant *Staphylococcus aureus* (MRSA) pandemics – presently a major public health concern (Oliveira et al., 2001, 2002; Lindsay, 2013; Otto, 2013).

Due to their high efficacy and low toxicity, β-lactams are the most widely used class of antibiotics (Shahid et al., 2009). They inhibit bacterial cell wall biosynthesis through irreversible binding to the traspeptidase domain of penicillin-binding proteins (PBPs) (Ghuysen, 1991, 1994).

Staphylococcus species have a broad distribution in nature and consist of large populations. They are common commensals of the skin and mucous membranes of humans and animals (Kloos, 1980, 1986, 1997) and are ubiquitously recovered from the environment (Huijbers et al., 2015). Although during most of its existence they live as mere colonizers, when the skin and mucous membranes barrier of their host is impaired and the host is immunocompromised staphylococci may arise as important pathogens. Among all staphylococcal species, *S. aureus*, is considered to be the most pathogenic, being associated to a myriad of infections ranging from mild skin infections to life-threatening diseases (Crossley and Archer, 1997).

The major driving force for the emergence of β-lactams resistance in staphylococci was the continuous exposure to β-lactams in multiple environments: in soils where they had to co-exist with penicillin-producing fungi; in production animal farms wherein large amounts of β-lactam antibiotics were used as food additives (National Research Council, 1980; Castanon, 2007), and during treatment of bacterial infections (Westh et al., 2004).

All MRSA contain a copy of an exogenous *mec* gene that codifies for PBPs with low affinity for β-lactams (*mecA*, *mecB*, *mecC*, and *mecD*) (Hartman and Tomasz, 1984; Harrison et al., 2013; Gomez-Sanz et al., 2015; Schwendener et al., 2017; Becker et al., 2018; Schwendener and Perreten, 2018), In this review we will focus on the evolution and emergence of methicillin resistance mediated by *mecA* which encodes an extra PBP, PBP2A (Hartman and Tomasz, 1984) with a low binding affinity to virtually all β-lactams. In the presence of β-lactams antibiotics the transpeptidase domain of all native PBPs is inactivated, but bacteria containing *mecA* continue to synthesize cell wall as a result of the cooperation between transpeptidase domain of the PBP2A and the transglycosylase domain of the native staphylococcal PBP2 (Pinho et al., 2001). The few β-lactam to which *mecA* does not confer resistance include ceftobiprole and ceftaroline which are active against MRSA (Entenza et al., 2002; Ishikawa et al., 2003) and penicillin G, ampicillin and amoxicillin

which are active against penicillinase-negative MRSA strains (a minority, nowadays).

Several efforts have been made to clarify the origin of *mecA*-mediated resistance to β-lactams in staphylococci and the use of state-of-the-art WGS technology has provided unprecedented advances (see **Table 1**). Nevertheless, the precise steps that led to β-lactam resistance development and dissemination are still not totally clear and are a matter of speculation.

HISTORY OF β-LACTAMS AND β-LACTAM RESISTANCE

Penicillin, a natural antibacterial compound produced by fungi, was first discovered in 1928 by Alexander Fleming (Fleming, 1929). However, due to low production yield, instability of the compound and problems in purification it was only later, in 1941, that penicillin was used as an antibiotic to treat human bacterial infections. The necessity to treat sick and wounded soldiers in the Second World War promoted the mass production of penicillin, and in 1945 this antibiotic was already used routinely in human clinical practice (Aminov, 2017).

Based on studies showing that penicillin is a growth promoter of chickens, pigs and livestock (National Research Council, 1980; Castanon, 2007), in 1951, the Food and Drug Administration (FDA) also approved the use of penicillin in animals (Hao et al., 2014). Nowadays, penicillin and other

---

TABLE 1 | Insights into β-lactam resistance development provided by WGS.

***mecA* evolution**

- *mecA* homologs are ubiquitous in *S. fleurettii* and *S. vitulinus*
- *mecA* homologue is present in the native location In *S. fleurettii* and *S. vitulinus*
- *mecA* and *mecA2* in *S. vitulinus* do not provide resistance to β-lactams
- *mecA* homologs diversification begun with the use of antibiotics (1940s)
- *rnecB* and *mecC* were Identified

***SCCmec* evolution**

- *mec* complex and SCC elements evolved in parallel in different chromosomal locations
- *mecA* homologs native location in *S. sciuri* species group is 200 Kb from *orfX*
- *mecA* and *mecR2* originated from *S. sciuri* homologs
- *mecRI/mecI* were added to *mecA* to form the *mec* complex in *S. fleurettii* or *S. vitulinus*
- The last donors of J1 region and *ccr* to SCC*mec* were *S. sciuri*
- The last donors of *mec* complex, J2 and J3 regions to SCC*mec* were *S. vitulinus/S. fleurettii*
- Origin of SCC and SCC*rnec* is probably *S. sciuri*
- *mecA* was probably integrated into a SCC in 5. *sciuri*
- SCC*mec* III was probably the most ancient SCC*mec* type

Expression of β-lactam resistance

- Genetic basis of hetero-to-homo resistance conversion
  - Tandem amplification of SCC*mec*
  - Mutations in *relA* and *rpoB*
- Mechanisms of β-lactam resistance development in *S. sciuri* and S. *vitulinus*
  - Alterations in *mecAI/mecA2* promoter
  - Alterations in PbpD structure
  - SCC*mec* acquisition

β-lactams continue to be used in many countries in food production animals not only to enhance animal growth, but also to treat infections and as a prophylactic (Hao et al., 2014). In fact, recent surveillance studies in Europe indicate that 25% of all antibiotic consumption (in mg/PCU) in veterinary setting relate to penicillin (report SE, 2010/2015), much of which are used for non-therapeutic purposes in chickens, cattle, and swine, compared with just a small quantity used for clinical treatments (World Organization for Animal Health, 2016).

A natural consequence of penicillin exposure was the development of antibiotic resistance. In fact, in 1942, only 2 years after the introduction of penicillin into clinical practice, the first penicillin-resistant *S. aureus* emerged in a hospital (Barber and Rozwadowska-Dowzenko, 1948) and shortly after (1960s) were also disseminated in the community (Rountree and Freeman, 1955), reaching around 80%. Penicillin resistance emerged due to the acquisition of β-lactamases that were able to hydrolyze and inactivate penicillin. Further developments to overcome resistance to penicillin included the synthesis of penicillinase-resistant penicillins, such as methicillin in 1960. However, due to its amazing adaptive power, *S. aureus* that were resistant to methicillin and to all β-lactams emerged right after its first use in the treatment of bacterial infections, through the acquisition of *mecA* (Jevons, 1961). This event has lead to the emergence of methicillin-resistant *S. aureus* (MRSA) strains and to one of the most important bacterial pandemics in hospitals worldwide (Oliveira et al., 2002; Grundmann et al., 2006). The MRSA rates in hospitals increased then, exponentially, reaching extremely high levels (above 60%) in the 1990s, mainly in Southern European countries (Deurenberg and Stobberingh, 2008; Chambers and Deleo, 2009; Lindsay, 2013). Similar to what was observed before for penicillin, this was followed by a wave of MRSA emergence in the community (Herold et al., 1998; Continuous Discharge Certificate [CDC], 1999; Vandenesch et al., 2003; Vlack et al., 2006), causing infections in otherwise healthy persons. Community-associated MRSA (CA-MRSA) are nowadays endemic in the community in specific countries like United States and here they have also become major multidrug resistant hospital clones (Moran et al., 2006; Otto, 2013). Additionally, resistance to β-lactams has expanded into farm environments wherein specific MRSA clones, like ST398 have become frequent colonizers of production animals and of humans in contact with them (Armand-Lefevre et al., 2005; Voss et al., 2005; Fluit, 2012).

As was observed for β-lactams, the emergence of drug resistance has been described following the introduction of each new antimicrobial class. The recent awareness by political authorities of the problem of antimicrobial resistance has lead to actions toward the banning of antimicrobial use in animals (European Union [EU], 2003; Food and Drug Administration [FDA], 2018). However, rules controlling antimicrobial use in animals have been applied mainly in Europe (National Research Council, 1980; World Organization for Animal Health, 2016), and still some of the antimicrobials, like penicillins, used to treat human disease continue to be heavily used in animals in European countries (Grave et al., 2012).

THE STRUCTURAL ELEMENT OF β-LACTAM RESISTANCE: THE *mecA* GENE

β-lactams target the PBPs, involved in the synthesis of peptidoglycan, the major structural component of the bacterial cell wall. In particular, PBPs catalyze the main reactions involved in the polymerization of peptidoglycan, namely the elongation of glycan strands (transglycosilation) and the cross-linking between stem peptides of different glycan strands (transpeptidation) (Macheboeuf et al., 2006). Binding of β-lactams to native PBPs turns them inactive, what prevents peptidoglycan synthesis and bacterial growth (Waxman and Strominger, 1983). This reaction involves the break of the β-lactam ring amide bond and acylation of the PBPs, which gives rise to a serine ester-linked acyl derivative that is extremely stable and has a low rate of deacylation.

The *mecA* gene encodes a high-molecular weight class B PBP, called PBP2A (Hartman and Tomasz, 1984), which contains two domains, the C-terminal domain which is known to have a transpeptidation function, and a N-terminal domain to which no function has been attributed, the so called non-binding (NB) domain. Resistance is provided by the fact that this extra PBP has a lower efficiency of acylation by β-lactams, which is believed to result from a lower affinity for these compounds (Fuda et al., 2004) and from a slow rate of acylation. Resolution of PBP2A crystal structure showed that the poor acylation rate observed is due to the presence of a distorted active site, provided by a higher flexibility of the NB domain and by differences in regions close to the active site groove in the transpeptidase domain (Lim and Strynadka, 2002). Furthermore, the position of Ser403 was considered crucial for the effective nucleophilic attack of the β-lactam ring, which leads to acylation of the protein (Lim and Strynadka, 2002).

THE MOBILE ELEMENT CARRYING *mecA*: SCC*mec*

The *mecA* gene is carried in a mobile genetic element called staphylococcal cassette chromosome *mec* (SCC*mec*) (International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements [IWG-SCC] et al., 2009). SCC*mec* is delimited by distinctive terminal inverted and direct repeats formed upon SCC*mec* insertion (DR-left downstream *orfX*, attL; DR-right at SCC*mec* end, attR) in a single chromosomal location, in the 3′ end of *orfX*, a RNA methyltransferase that is localized near the origin of replication (Boundy et al., 2013). This mobile genetic element is composed of two central elements, the *mec* complex containing *mecA* and intact and deleted forms of its regulators (*mecI*, *mecR1*) and the *ccr* complex composed of cassette chromosme recombinases (*ccr*) involved in its mobility (Katayama et al., 2000). The remaining portions of SCC*mec*, are composed of non-essential components, namely additional metal and antibiotic resistance

genes carried by transposons and plasmids, as well as genes of unknown function, which are named J regions. The J3 region is located between *orfX* and *mec* complex, the J2 region is flanked by *mec* complex and the *ccr* complex and the J1 region between *ccr* complex and the right extremity of the element (see **Figure 1**). More recently, *mecR2*, coding for an anti-repressor of *mecA* was described to exist downstream *mecI*, which together with *mecI* and *mecR1* constitute an unusual three-component arrangement (Arede et al., 2012). So far as many as thirteen different structural types of SCC*mec* have been described in *S. aureus*[1] (Ito et al., 2001, 2004; Ma et al., 2002; Oliveira et al., 2006; Berglund et al., 2008; Zhang et al., 2009; Garcia-Alvarez et al., 2011; Li et al., 2011; Wu et al., 2015; Baig et al., 2018) that range between 20 and 70 Kb. The different types of SCC*mec* correspond to different combinations of *mec* complex class (A-E), according to the presence/absence of regulatory genes and insertion sequences, and *ccr* allotypes (*ccrAB* and *ccrC*).

SCC*mec* is believed to have been acquired a limited number of times in *S. aureus* (Robinson and Enright, 2003), but acquisition of this element appears to provide a major advantage to bacteria mainly in the hospital environment. In fact, the acquisition of different types of SCC*mec* by methicillin-susceptible *S. aureus* (MSSA) of diverse genetic backgrounds gave rise to several MRSA pandemics over time (Chambers and Deleo, 2009), namely the Iberian (ST247-I), Brazilian (ST239-III), New York-Japan (ST5-II), EMRSA16 (ST36-II), EMRSA15 (ST22-IV), Berlin (ST45-IV), United States 300 (ST8-IV), and ST398-V (Deurenberg and Stobberingh, 2009). In contrast to *S. aureus*, in coagulase-negative staphylococci (CoNS), there is not a clear association between SCC*mec* and specific genetic backgrounds (Rolo et al., 2012). However, it is still not clear if this derives from the fact that SCC*mec* was acquired a higher number of times or from a higher instability of SCC*mec* structure in these species. Nevertheless, specific SCC*mec* types appear to be more common to certain CoNS species. SCC*mec* type I and VI are more common in *Staphylococcus hominis* (Bouchami et al., 2012), SCC*mec* III in *Staphylococcus sciuri* group of species (Rolo et al., 2017a), SCC*mec* IV in *S. epidermidis* (Miragaia et al., 2007), and SCC*mec* V in *S. haemolyticus* (Bouchami et al., 2010).

The true breadth of genetic diversity of SCC*mec* elements is still unknown. While in *S. aureus* this structure has been shown to be relatively stable, in CoNS, such as *S. epidermidis* and *S. haemolyticus*, a high genetic diversity has been described (Miragaia et al., 2007). This might be due to the fact that CoNS species have a high recombination rate (Miragaia et al., 2007), an enhanced ability to acquire and maintain exogenous genetic material or because these elements have been acquired earlier by these species than by *S. aureus*. In the era of WGS where detailed information on the entire genomes of thousands of staphylococci are being gathered, the number of types and subtypes of SCC*mec* have increased exponentially[1], and this challenges the traditional criteria and methodologies that were previously defined to classify SCC*mec* types, mainly based on

PCR. A web-tool, SCC*mec*Finder, able to identify all SCC*mec* element types (I to XIII) and SCC*mec* IV and V subtypes was recently developed to classify SCC*mec* types based on WGS data (Kaya et al., 2018). The characterization of the SCC*mec* elements is based on two different gene prediction approaches to achieve correct annotation.

THE SCC ELEMENTS

Although due to its clinical relevance, SCC*mec* is the most popular element, several other SCC elements not carrying the *mec* complex (SCC) or either *mec* complex and *ccr* complex (pseudo-SCC) have been identified at the *orfX* site. These elements can carry diverse genes relevant for staphylococcal survival and virulence, namely conferring heavy metal resistance genes (Chongtrakool et al., 2006) providing capsule production (Luong et al., 2002), cell-wall biosynthesis (Mongkolrattanothai et al., 2004), restriction/modification functions or immune protection (Holden et al., 2004). These elements can be found alone in the chromosome or in tandem with SCC*mec* or other SCC elements, being in this case named composite islands (CI). Examples of such CIs include the SCC*mec* III-SCC*mer* from the pandemic MRSA Brazilian clone (Chongtrakool et al., 2006) and the SCC*mec* IV-ACME from the USA300 CA-MRSA clone (Shore et al., 2011b). Although many studies have described the structure and contents of these elements, few studies have addressed their true contribution for staphylococcal virulence or fitness (Diep et al., 2008).

Besides being inserted at the exact same location as SCC*mec*, SCC and pseudo-SCC elements have been described to have regions of homology with SCC*mec* (Katayama et al., 2003a) (see **Table 2**), suggesting that their evolutionary history is related. Still, until recently the nature of their relatedness remained elusive.

TRANSFER OF SCC ELEMENTS

SCC*mec* with exactly the same nucleotide sequence were found in different staphylococcal strains and species, suggesting that this element is frequently transferred among staphylococci (Shore et al., 2011b). However, the mechanisms of SCC and SCC*mec* insertion and excision from the chromosome as well as the mechanism of transfer is still elusive and it is likely that several different mechanisms are involved.

When grown in antibiotic free medium SCC*mec* can be excised from the chromosome, a reaction catalyzed by the cassette chromosome recombinases (Ccr) (Katayama et al., 2000). Immediately after excision, SCC*mec* circularizes into an extrachromosomal intermediate form and the attSCC sites are created in the chromosome (*attB*) and the intermediate circular form (*attS*). The created attSCC sites, work like recognition sites for the Ccr enzymes for a later integration into the chromosome. During SCC*mec* excision and insertion *orfX* remains always intact (Katayama et al., 2000).

Electrophoretic mobility shift assay (EMSA) showed that Ccr enzymes recognize a minimum of 14-bp sequence in the

---

[1]http://www.sccmec.org/Pages/SCC_TypesEN.html
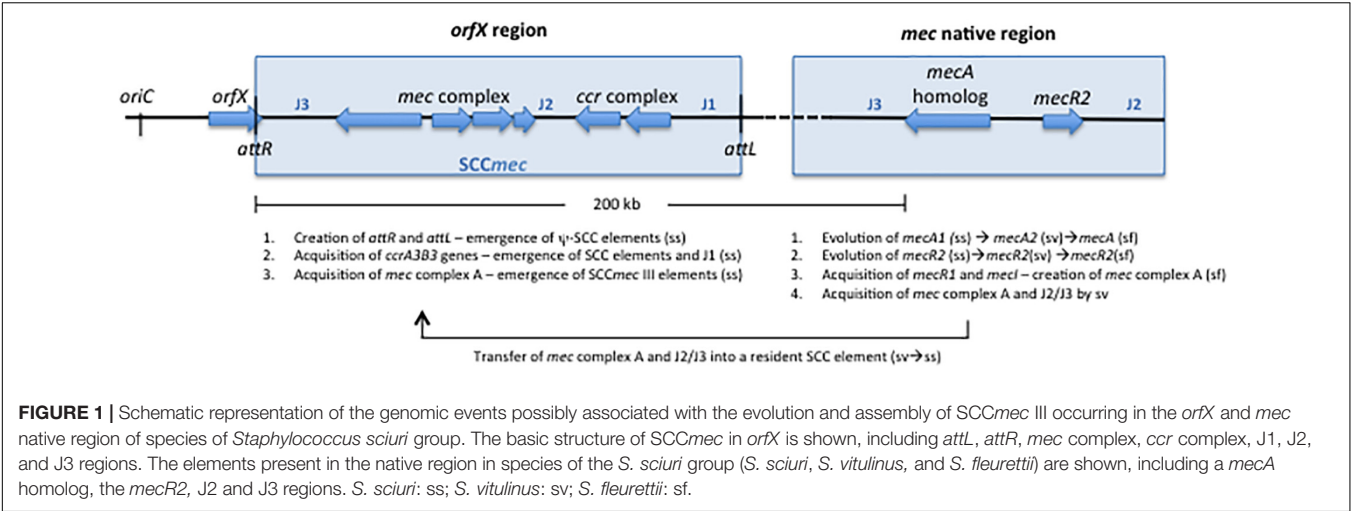
FIGURE 1 | Schematic representation of the genomic events possibly associated with the evolution and assembly of SCC*mec* III occurring in the *orfX* and *mec* native region of species of *Staphylococcus sciuri* group. The basic structure of SCC*mec* in *orfX* is shown, including *attL*, *attR*, *mec* complex, *ccr* complex, J1, J2, and J3 regions. The elements present in the native region in species of the *S. sciuri* group (*S. sciuri*, *S. vitulinus,* and *S. fleurettii*) are shown, including a *mecA* homolog, the *mecR2*, J2 and J3 regions. *S. sciuri*: ss; *S. vitulinus*: sv; *S. fleurettii*: sf.

TABLE 2 | Nucleotide identity (%) of SCC elements found *in S. sciuri, S. vitulinus* and *S. fleurettii* with S. *aureus* SCC*mec*.

SCC element	Species	*ccr* type	*mec* complex	Homology with *S. aureus* SCC*mec* region (%)
SCC$_{11/01}$	S. sciuri	ccrA5B5	–	J1: IVa (46), SCC*mec* S. xylosus (mecC) (59)
				J2: III (67–96), IX (79–93)
				J3: IX (69–91)
SCC$_{K6-15937}$	S. sciuri	ccrA5Bnew	–	J1:III(90), V(79)
				J2: III (95)
				J3: IX (99)
ψ-SCC$_{K6-10930}$	S. sciuri		–	Jl: V (65)
				J3: III (55–80)
SCC*mec* III-like B	S. sciuri	ccrA3B5	A	J2: II (99–100), III (60–100). IX (73)
				J3: III (100)
SCC*mec* III-like A	S. sciuri	ccrA3B5	A	J1: X (96), III (90–100)
				J2: III (38–100)
				J3: IV (42)
SCC$_{K116}$	S. sciuri	ccrC	–	J2: II (90), III (91–98), V (73–80)
SCC-Cl$_{SS27}$-I SCC-Cl$_{SS27}$ - II	S. sciuri	ccrA1B3 ccrA1B3		J2: III(77), X(90)
				Jl: IX (77), XI (73)
				J2: II (94), IV (62–79), IX (73–97)
SCC$_{402567}$	S. fleuretti	ccrAnewB5	–	J2: IX (75), III (77–96)
SCC$_{11683}$	S. vitulinus	ccrABnew	–	J1: IX (56–95) J2: IX (97)
SCC$_{401946}$	S. vitulinus	ccrABnew	–	J1: IX (76), V (82) J3: IV (51)

terminal sequence of the *orfX* that contains always the conserved sequence (TATCATAA), which is also found in SCC*mec* (*attS*). However, DNA sequences flanking the *att* sites were shown to also have a role on the frequency and efficiency of SCC*mec* insertion (Wang et al., 2012). This sequence specificity was thus probably important in the epidemiology of SCC*mec* acquisition by staphylococci and might partly explain the association of some SCC*mec* to specific genetic backgrounds or staphylococcal species (Oliveira et al., 2002; Miragaia et al., 2007; Bouchami et al., 2010, 2012; Rolo et al., 2017a).

According to studies of translational fusions of the *ccr* promoter with green fluorescent protein the *ccr* activity and associated SCC*mec* excision is a bistable process occurring only in a small fraction of cells within the population (Stojanov et al., 2013). The fate of the extrachromosomal circularized SCC elements after they are formed is still a mystery. SCC elements have once been regarded as non-replicative, due to the absence of a replication origin, but recent crystallographic studies have provided evidence that SCC*mec* elements encode an active MCM-like helicase (Mir-Sanchis et al., 2016), suggesting that they can eventually replicate in the cytoplasm before being transferred, but further studies are needed to confirm these observations.

The transfer of SCC*mec* have been successfully achieved in the laboratory by several different genetic mechanisms. In particular susceptible strains in contact with phage lysates containing this mobile genetic element were shown to become resistant to β-lactams and SCC*mec* elements were successfully packaged into

bacteriophage capsides (Cohen and Sweeney, 1970; Maslanova et al., 2013; Scharn et al., 2013; Chlebowicz et al., 2014; Haaber et al., 2016). Additionally, transfer of SCC*mec* was also achieved through natural transformation, upon induction of SigH in very specific laboratory growth conditions (Morikawa et al., 2012). Moreover, a chromosomally encoded and laboratory-constructed derivative of SCC*mec* was captured on a conjugative plasmid and transferred by filter-mating into different *S. aureus* and *S. epidermidis* recipients (Ray et al., 2016). Also, evidence of possible transfer of SCC*mec* by conjugation was the finding of a *mecA* homologue within a plasmid of *Macrococcus caseolyticus*, a species phylogenetically related to *Staphylococcus* (Tsubakishita et al., 2010a).

In spite of the huge effort of the scientific community in elucidating the mechanism of SCC*mec* transfer, many of the studies described occurred in very artificial conditions and it remains to be clarified which mechanism(s) are actually more frequent *in vivo*.

THE ORIGIN OF THE METHICILLIN RESISTANCE DETERMINANT – *mecA* EVOLUTION IN *S. sciuri* GROUP

Early studies, based on structural nucleotide identity, have proposed that the *mecA* gene originated from recombination between a PBP from *E. coli* (Song et al., 1987) or from *Enterococcus hirae* (Archer and Niemeyer, 1994) with a β-lactamase encoding gene. A theory that was later supported by the finding by WGS of another *mec* allotype (*mecC*) as part of a class E *mec* complex, containing *blaZ* (*mecI-mecR1-mecC-blaZ*) in *Macrococcus caseolyticus* (Tsubakishita et al., 2010a) and *S. xylosus* (Harrison et al., 2013).

Other lines of evidence suggest that *mecA* originated from native PBPs of species of the *Staphylococcus sciuri* group – a primordial phylogenetic clade, including *S. sciuri*, *Staphylococcus fleurettii*, *Staphylococcus vitulinus*, *Staphylococcus lentus*, and *Staphylococcus stepanovicii* (Schleifer et al., 1983; Couto et al., 1996; Zhou et al., 2008; Antignac and Tomasz, 2009; Hauschild et al., 2010; Tsubakishita et al., 2010b), which most important ecological niches are the soil and skin and mucous membranes of wild and production animals.

The use of WGS on a large collection of isolates belonging to the *S. sciuri* group revealed the presence of homologues of *S. aureus mecA* with different levels of homology that were ubiquitous within some of the species (Couto et al., 1996; Wu et al., 1996; Hiramatsu et al., 2013) (see **Figure 2A**). However, in contrast to *mecA* in *S. aureus*, *mecA* homologs in *S. sciuri* group species were all located approximately 200 Kb apart from *orfX* (native location) between *mva* and *xyl* operons, outside any SCC element (Rolo et al., 2017b) (see **Figure 2B**). The results suggest that *mecA* has been transmitted vertically during the early stages of staphylococcal speciation (see **Figure 2A**; Step 1, **Figure 3**).
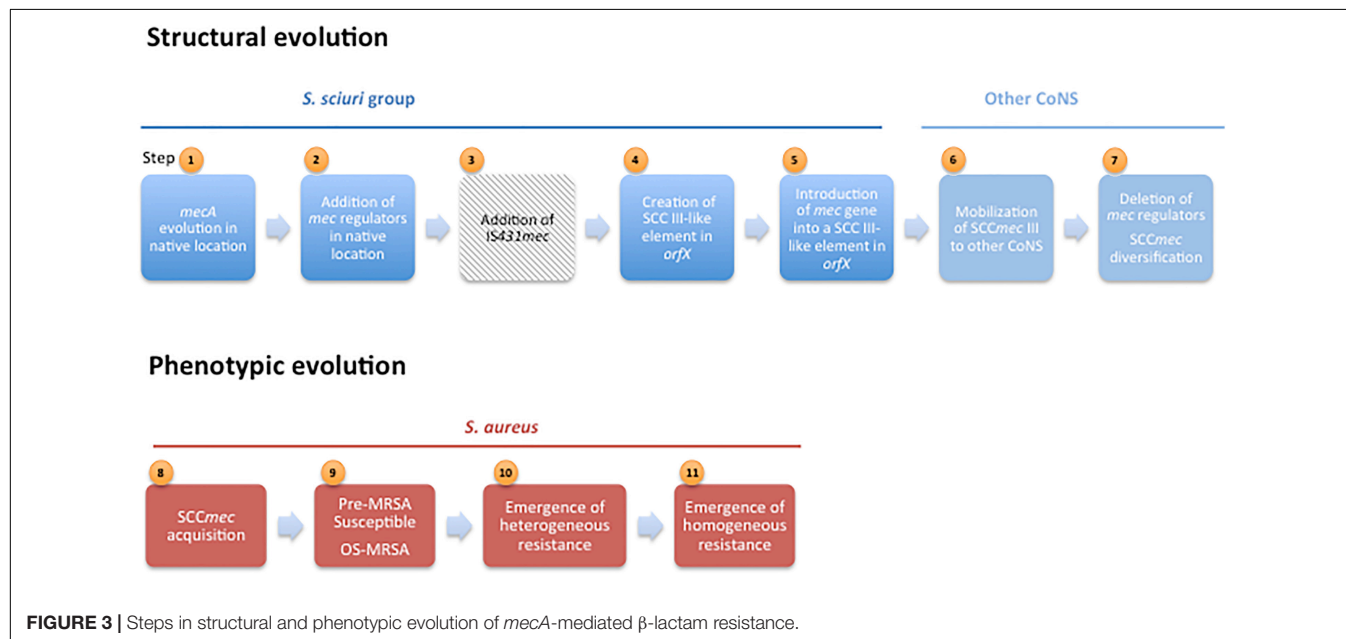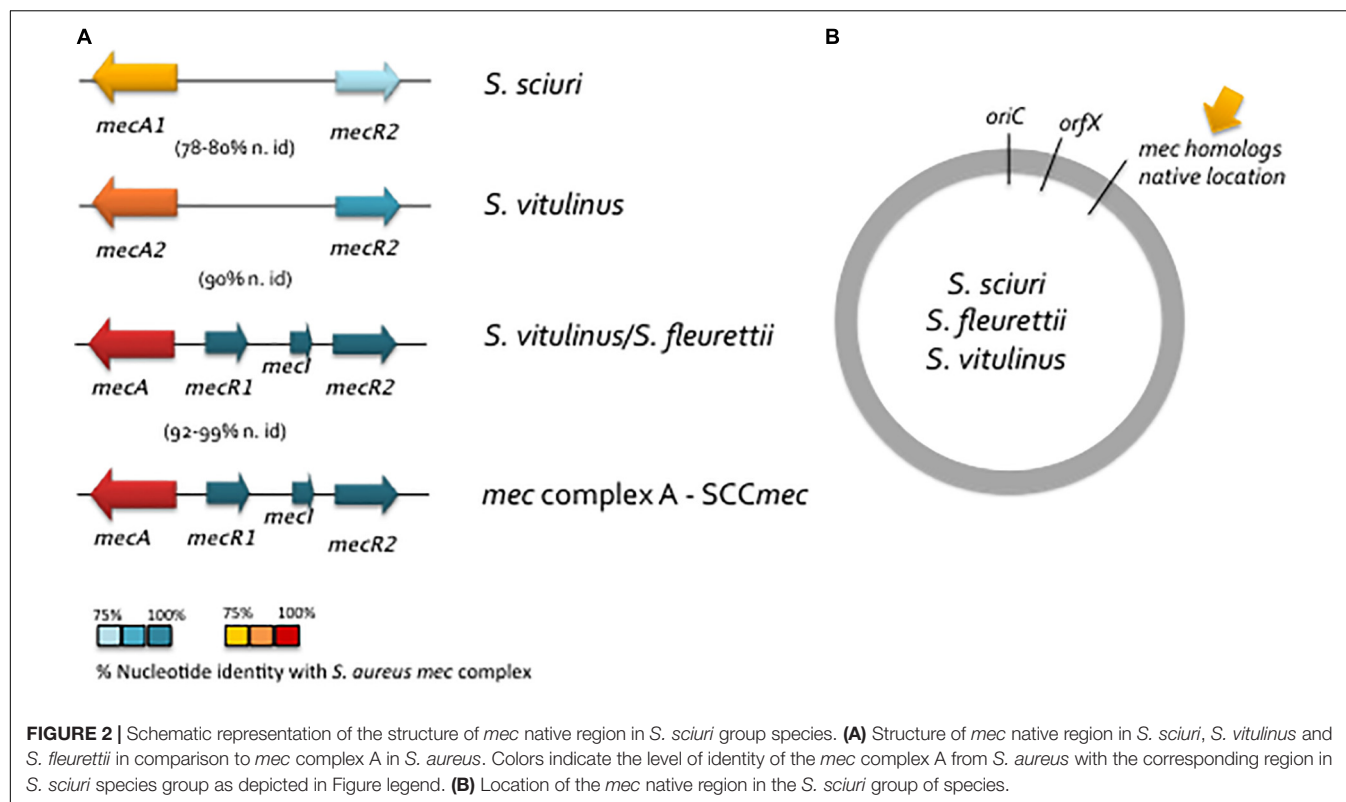
Although the primary function of these *mecA* precursors was probably related to cell wall synthesis and not to antimicrobial resistance, the complete evolution from a native PBP into a resistance determinant appears to have been a stepwise process

that occurred within this group of species. *S. sciuri* carries the most ancestral form of *mecA* (*mecA1*) which has 85% homology in nucleotide sequence with *S. aureus mecA*; *S. vitulinus* harbors an intermediary form (*mecA2*) with 94% homology and all *S. fleurettii* and some *S. vitulinus* have a *mecA* form that is almost identical to that of *S. aureus mecA* (*mecAf*, *mecAv*; 99% homology) (Rolo et al., 2017b; Tsubakishita et al., 2010b) (see **Figure 2** and **Table 3**). In *S. lentus* and *S. stepanovicii* so far no *mecA* homolog has been described, (Tsubakishita et al., 2010b; Calazans-Silva et al., 2014), but no extensive and detailed study was ever performed in these two species.

The only *mecA* homologue that confers resistance to β-lactams is the *mecA* in *S. fleuretti*. The same *mecA* homologue in *S. vitulinus* and the *mecA1* and *mecA2* found in *S. sciuri* and *S. vitulinus*, respectively, do not confer resistance to β-lactams in the great majority of strains (Couto et al., 1996; Wu et al., 1998; Monecke et al., 2012; Rolo et al., 2017b). However, *S. sciuri* and *S. vitulinus* strains exhibiting β-lactam resistance have been reported (Couto et al., 2003; Tsubakishita et al., 2010b; Rolo et al., 2017b). Recent work by Rolo et al. (2017b) wherein a large collection of *S. sciuri* and *S. vitulinus* were analyzed by WGS showed that β-lactam resistance in these species emerged multiple times during evolution and was driven mainly by the contact with human created environments, namely with the beginning of the use of antibiotics in production animals and humans. The mechanisms of resistance development in these two species included: (i) the structural diversification of the non-binding domain of native PBPs which altered the structure of the active site and exposure of ser403; (ii) mutations and insertion of IS*256* in the promoters of *mecA* homologs that were associated to an increased expression of the proteins; (iii) acquisition of SCC*mec*. Additionally, like in *S. aureus*, the bacterial genetic background plays an important role in the expression of β-lactam resistance in the *S. sciuri* group of species, since the exact same gene allele was found associated to both susceptible and resistant strains (Rolo et al., 2017b).

Additional evidence supporting that *S. sciuri mecA1* was the evolutionary precursor of *S. aureus mecA*, include the fact that this gene could be recruited to express methicillin resistance in *S. sciuri* after stepwise exposure to methicillin (Wu et al., 2001). Moreover, the activated copy of *S. sciuri mecA1* was, similarly, able to restore methicillin resistance phenotype, when transduced into methicillin-susceptible *S. aureus* (MSSA), conferring high level, homogeneous and broad-spectrum β-lactam resistance (Wu et al., 2001). Furthermore, the *S. sciuri mecA1* when transduced into MSSA was shown to act exactly like *S. aureus mecA*, being controlled by *S. aureus mecA* regulators (*mecI* and *mecR1*) and its product (PBP4) taking part in cell wall biosynthesis, producing a peptidoglycan typical of methicillin-resistant *S. aureus* (Antignac and Tomasz, 2009).

Besides *mecA*, other *mec* genes have been identified that are associated with β-lactam resistance, namely *mecB* and *mecD* in *Macrococcus caseolyticus* (Baba et al., 2009; Schwendener et al., 2017; Schwendener and Perreten, 2018) and *mecC* in *S. aureus* (Garcia-Alvarez et al., 2011; Shore et al.,

FIGURE 2 | Schematic representation of the structure of *mec* native region in *S. sciuri* group species. **(A)** Structure of *mec* native region in *S. sciuri*, *S. vitulinus* and *S. fleurettii* in comparison to *mec* complex A in *S. aureus*. Colors indicate the level of identity of the *mec* complex A from *S. aureus* with the corresponding region in *S. sciuri* species group as depicted in Figure legend. **(B)** Location of the *mec* native region in the *S. sciuri* group of species.

FIGURE 3 | Steps in structural and phenotypic evolution of *mecA*-mediated β-lactam resistance.

2011a), *S. xylosus* (Harrison et al., 2013), *S. sciuri carnaticus* (Harrison et al., 2014), and *S. stepanovicii* (Loncaric et al., 2013). The *mecB* and *mecD* are the most distant from *S. aureus mecA*, having, respectively, a nucleotide identity with *mecA* that is equal or lower than 62%, whereas *mecC* has 69% nucleotide sequence identity. All *mec* forms confer resistance to β-lactams to their natural hosts and their introduction

into a susceptible *S. aureus* genetic background was able to provide a resistance phenotype, confirming that they should encode a PBP with low-affinity to β-lactams that participates in cell wall synthesis (Baba et al., 2009; Kim et al., 2012). Both *mecB* and *mecC* were carried within mobile genetic elements structurally similar to SCC*mec* that were inserted in the *orfX* region (SCC*mec* XI, SCC*mec*

TABLE 3 | Nucleotide identity (%) of chromosoomal regions of *S. sciuri, S. fleurettii* and *S. vitulinus* with those found within *S. aureus* SCC*mec*.

SCC*mec* clement	Nucleotide identity with *S. aureus* SCC*mec* element (%)		
	S. sciuri	*S. vitulinus*	*S. fleurettii*
J3 REGION			
ugpQ	75–85	87–100	99–100
maoC	100	100	100
pre	64	65	63
polypeptide B	63	99	99
mec complex			
IS431	99	99	99
IS1272	63	66	63
mecA1	80	–	–
mecA2	–	90	–
mecA	–	99–100	97–99
mecRI/mecI	–	99/92-99	99/92
mecR2	78	99	99
J2 REGION			
psm-mec	91–94	97–99	98–99
Rhodanese-domain containing protein	77	100	99
Metalo- betalactamase family protein	78	99	99
ccr complex			
ccr	72–100	72–77	83
multiple *ccr*	–	–	–
J1 REGION			
Hypothetical proteins	90–100	90–100	17

For each chromosomal region, yellow cells represent the lowest %identity, red cells represent the highest %identity and orange cells the intermediate %identity.

IX-like) (Gomez-Sanz et al., 2015) and *mecB* was additionally found within a plasmid in *M. caseolyticus* (Baba et al., 2009; Tsubakishita et al., 2010a). The *mecD* gene is carried within a resistance island (McRI_{mecD}-1, McRI_{mecD}-2) that is inserted in 3′ end of the *rpsI* gene. Besides *mecD* this island contains genes for an integrase of the tyrosine recombinase family, but does not resemble either SCC elements or *mecB*-carrying mobile genetic elements (Schwendener et al., 2017; Schwendener and Perreten, 2018). However, none of the *mecB, mecC* or *mecD* was found within the native location (200 Kb apart from *orfX*).

The exact evolutionary link between *mecA, mecB, mecC* and *mecD* forms is still undetermined. Among all *mec* genes, *mecA* is apparently, the most successful in *Staphylococcus*. The *mecB* was recently found within a plasmid in a single *S. aureus* human carriage strain belonging to ST7 (Becker et al., 2018) and *mecC* has been limited to only a few *S. aureus* clonal lineages (CC130 and ST425) and four *Staphylococcus* species (*S. sciuri, S. xylosus, S. stepanovicci*, and *S. aureus*) (Harrison et al., 2013, 2014; Loncaric et al., 2013; Becker et al., 2014; Semmler et al., 2016). MRSA harboring *mecC* are believed to have a zoonotic origin and although they were reported in several different countries, they have been rarely observed in human infection (Becker et al., 2014). However, surveillance of dissemination of these *mec* genes should not be disregarded, since antibiotic use and the consequent selective pressure could drive fast evolutionary leaps that can lead to their precipitous spread.

STAGES IN THE EVOLUTION OF SCC*mec*

Most of the efforts have been focused on the clarification of the origin and evolution of the β-lactam resistance determinant (*mecA*). Much less information is available regarding the evolution of SCC*mec*, the mobile element carrying *mecA*, which is responsible for the worldwide spreading of β-lactam resistance among staphylococci. SCC*mec* is a mosaic-like element that was described to contain multiple transposable elements, plasmids and insertion sequences in J regions (Ito et al., 2001), a genetic environment that *per si* can promote and facilitate genetic variation and recombination, what has been hindering the reliable tracing of their phylogeny.

The characterization of the native location of *mecA* homologs, the SCC insertion site and the genetic background of a large collection of isolates belonging to *S. sciuri* group by comparative genomics showed that SCC elements and *mecA* and flanking regions evolved in parallel in these species in these two distinct chromosomal locations (Rolo et al., 2017b).

Assembly of the *mec* Complex in the Native Location

The *mecA* homologs flanking genes in the native location were found to be the same as those flanking *mecA* inside SCC*mec*, encompassing the J2 and J3 regions (Rolo et al.,

2017b) (see **Table 3**). Moreover, as for native *mecA* homologs, the level of homology of their flanking genes (*psm-mec* and *ugpQ*) with the same genes in SCC*mec* from MRSA, varied according to the phylogeny, wherein those of *S. fleuretti* were the most similar and those of *S. sciuri* were the most distant (Rolo et al., 2017b). The results suggest that the first stage of SCC*mec* evolution included the evolution of *mecA* homologs and their neighbor genes in the native location (Step 1, **Figure 3**). This was followed by the creation of the *mec* complex (Step 2, **Figure 3**). The *mecR2* was the first regulator to be added in *S. sciuri* at the native location near *mecA1*, since the most ancient precursor of *mecR2* was found in this species. This gene organization was preserved along phylogeny and became ubiquitous in *S. fleurettii* and *S. vitulinus* (Tsubakishita et al., 2010b). Addition of *mecR1* and *mecI* happened later, after the evolution of the ancestral *mecA1* into *mecA* was complete, as demonstrated by the lack of these regulators in *S. sciuri* and their occurrence in the native location of *S. fleurettii* and *S. vitulinus* near *mecA* (Tsubakishita et al., 2010b; Rolo et al., 2017b). These finding came to reconcile previous controversies, suggesting that although *mecA1* was the original precursor of *mecA*, *S. fleuretti*/*S. vitulinus* were probably the last donors of the *mec* complex to give rise to SCC*mec*. The addition of IS*431* element probably occurred later, after *mecA*, regulators and neighboring regions were mobilized into a SCC element located in the *orfX* region (Step 3, **Figure 3**). Alternatively, it could have been added during their mobilization, as it was never detected in the native location in any of the strains tested.

Acquisition and expression of *mecA* in species in which this gene is not native imposes a fitness cost to bacteria (Ender et al., 2004). For this reason, the step of addition of regulators with subsequent *mecA* repression appears to have been particularly crucial in the maintenance of the gene in new host species (Katayama et al., 2003b) and thus in *mecA* dissemination. In fact, some of the first methicillin resistant staphylococci, the so-called pre-MRSA (Step 9, **Figure 3**) (Hiramatsu et al., 1992; Kuwahara-Arai et al., 1996), contained intact regulators and a susceptible phenotype.

Studies wherein the *mec* complex region was characterized in MRSA and MR-CoNS revealed that although *mecI* and the 3′ end of *mecR1* are deleted in a great proportion of contemporary clinical strains (Katayama et al., 2001), the 5′ end portion of *mecR1* as well as a copy of IS*431* downstream *mecA* (IS*431*-R) are conserved in every strain. Moreover, deletion of *mecI* and *mecR1* promoted by IS*431* was accomplished *in vitro* upon selection with methicillin in a *S. haemolyticus* with a *mec* complex type A (Suzuki et al., 1993). These observations are in accordance with the view that the *mec* regulators and the IS*431*-R together with *mecA* were once the original components of the *mec* region DNA and that deletion of the regulators occurred at a later time in evolution (Step 7, **Figure 3**). On the other hand, the similarity of nucleotide sequence in regions located upstream of the *mec* complex in four different *mec* complex classes, suggests that deletion of the *mec* regulators must have occurred after the establishment of the prototypic

mec complex A in a SCC*mec* element (Katayama et al., 2001).

Thus the assembly of *mec* complex A seems to be the first step of genetic evolution, followed by its establishment in SCC*mec* and subsequent deletion of the regulators to originate the different *mec* complex classes (*mec* complex B, C, D, E).

SCC Element Evolution in the *orfX* Region

Analysis of the *S. sciuri orfX* region showed that SCC elements most probably originated in *S. sciuri* and were assembled from housekeeping genes located in this region (Step 4, **Figure 3**), as evidenced by the finding of the same housekeeping genes either outside or within SCC elements (Rolo et al., 2017a) – a finding not observed in the other species of the *S. sciuri* group. Additionally, in *S. sciuri* the most ancestral forms of cassette chromosome recombinases (*ccr*) and the highest genetic diversity were found, including almost all *ccr* allotypes described in *S. aureus* (Rolo et al., 2014). Interestingly, it was also in this species that a SCC*mec* type III-like was found with high homology simultaneously with the *mec* complex and J2 region of *S. aureus* SCC*mec* type III and the J1 region and an ancestral form of *ccrAB3* of *S. sciuri* SCC non-*mec* (Rolo et al., 2017a) (see **Table 2**). The results suggest that SCC*mec* III originated in *S. sciuri,* probably through the integration of the *mec* complex and J2 region from *S. vitulinus*/*S. fleuretti* into a resident SCC non-*mec* carrying *ccrAB3* (Step 5, **Figure 3**). However, the mechanism that mobilized the *mec* complex from *S. vitulinus*/*S. fleuretti* to an SCC in *S. sciuri* is still not known. Once formed SCC*mec* III probably disseminated to other CoNS species, namely *S. epidermidis* and *S. aureus* ST239, wherein SCC*mec* III was found to be prevalent (Miragaia et al., 2007; Harris et al., 2012) (Step 6, **Figure 3**).

Although it is apparent that the origin of SCC*mec* type III is *S. sciuri*, the source of the remaining SCC*mec* types elements remains unclear. In contrast to *S. sciuri*, which carried a high diversity of *ccr* allotypes, methicillin susceptible CoNS species belonging to more recent clades in the phylogeny of staphylococci, which include *Staphylococcus epidermidis*, *Staphylococcus haemolyticus* and *Staphylococcus hominis*, were particularly enriched in a specific allotype of *ccr*. The *ccrAB2* was found to be common in *S. epidermidis* (Miragaia et al., 2007), *ccrAB1* in *S. hominis* (Bouchami et al., 2012) and *ccrC* in *S. haemolyticus* (Bouchami et al., 2010), which is coincidently the same type of *ccr* carried by the most frequent SCC*mec* in these species. It is thus tempting to speculate that each SCC*mec* type can result from the integration of *mec* complex, probably through recombination, into a resident SCC element in these species. This hypothesis is supported by the identification of SCC non-*mec* elements carrying different *ccrAB* types in *S. aureus* and CoNS with regions of high homology with known SCC*mec* types (Katayama et al., 2003a). However, the enrichment of certain SCC*mec* types in particular species may also derive from the described specificity of the different types of Ccr enzymes as described above (Wang et al., 2012).

SCC*mec* Diversification

The next stage of SCC*mec* evolution that is believed to be still ongoing includes the diversification and dissemination of the SCC*mec* element among the staphylococcal population (Step 7, **Figure 3**). The existence of similar regions among different SCC*mec* types, like the J1 region in SCC*mec* type II and IV or the *mec* complex B between SCC*mec* type I and IV (Chongtrakool et al., 2006) suggest that the different SCC*mec* types are related.

The involvement in the diversification process of species, such as *S. epidermidis*, *S. hominis* and *S. haemolyticus* is apparent. Besides being reservoirs of specific types of SCC*mec* and *ccr* allotypes, these species harbor a huge number of non-described SCC*mec* types (Wisplinghoff et al., 2003; Miragaia et al., 2007; Bouchami et al., 2010), evidencing their key role in the current diversification of SCC*mec*.

Moreover, factors associated to hospital environment appear to be driving the diversification and acquisition of SCC*mec* in species like *S. epidermidis* (Rolo et al., 2012). One of the factors related to the clinical setting that might be triggering SCC*mec* diversification is the use of antibiotics, namely β-lactams and vancomycin, which were already shown to promote the expression of recombinases (Higgins et al., 2009). The excision of SCC elements promoted by *ccr* overexpression may create new opportunities of recombination between different elements within the same strain, giving rise to new SCC*mec* structures.

THE IMPORTANCE OF GENETIC BACKGROUND FOR THE ACQUISITION OF SCC*mec* AND FOR THE EXPRESSION OF β-LACTAM RESISTANCE IN *S. aureus*

The transfer of SCC*mec* from CoNS to *S. aureus* was probably a subsequent step (Step 8). Several lines of evidence suggest that acquisition and expression of *mecA* by *S. aureus* was a complex process involving multiple genetic and metabolic alterations. The construction of a Tn*551* transposon library in the background of the MRSA strain COL and subsequent screening for a decreased level of methicillin resistance has identified several factors (fem or auxiliary genes) that, together with the *mecA* gene, are crucial for the expression of high-level and homogeneous resistance to methicillin (de Lencastre and Tomasz, 1994) (Step 11, see **Figure 3**). Although having a substantial impact on oxacillin resistance, these genes are not directly implicated in the expression of *mecA*, but they are mainly involved in cell wall metabolism and stress response (de Lencastre and Tomasz, 1994). However, these studies were performed in a single MRSA strain (COL) and these same genes appear to have different contributions to β-lactams resistance in other *S. aureus* genetic backgrounds (Memmi et al., 2008; Figueiredo et al., 2014), suggesting that expression of β-lactam resistance is extremely complex and that auxiliary genes in different MRSA strains might be different or use different mechanisms. Additionally, whether the identified auxiliary genes in COL also contribute to β-lactam resistance expression in other *Staphylococcus* species is unknown.

What appears to hold true is that not all *S. aureus* appear to have the same ability to accommodate *mecA*. The existence of a host barrier was evidenced by the finding that SCC*mec* was acquired by a limited number of *S. aureus* genetic backgrounds (e.g., ST239, ST45, ST22, ST8, ST5) (Robinson and Enright, 2003) while other genetic backgrounds despite being successful, like ST121, were rarely observed carrying *mecA* (Rao et al., 2015). Furthermore, when a recombinant plasmid, carrying intact *mecA*, was introduced into strains that have never experienced the presence of *mecA,* they were unable to maintain or express β-lactam resistance, a phenomena not observed when the same assay was performed in MSSA strains from which SCC*mec* has been excised (Katayama et al., 2003b). Interestingly, either the presence of β-lactamase (*blaR1-blaI*) or *mecA* regulatory genes (*mecR1-mecI*), which control *mecA* expression, allowed the maintenance and expression of plasmid-carried *mecA* in the naïve genetic background (Katayama et al., 2003b), which is indicative that besides the genetic backgrounds the repression of *mecA* was important for the acquisition and stability of *mecA* in staphylococci. Actually, the so-called pre-MRSA although carrying *mecA*, showed a susceptible phenotype, which was shown to result from *mecI*-mediated repression of *mecA* transcription (Kuwahara-Arai et al., 1996) (Step 9, **Figure 3**). The integration of newly acquired genes into the recipient metabolic network is a complex mechanism that frequently represents a large fitness cost for bacteria. The presence of the regulators will probably work as safeguard mechanisms that will silence the newly acquired gene and avoid potentially harmful consequences of its expression in the new bacterial host, while it is still not adapted (Ochman et al., 2000; Navarre et al., 2006).

A different phenomenon supporting the importance of genetic background for the expression of β-lactam resistance is the emergence of the so-called oxacillin susceptible MRSA (OS-MRSA) (Step 9, **Figure 3**), strains that like pre-MRSA carry *mecA* and do not express β-lactam resistance, but that in contrast do not carry *mecI* (SCC*mec* IV or V) (Giannouli et al., 2010; Andrade-Figueiredo and Leal-Balbino, 2016; Phaku et al., 2016). OS-MRSA have been recently described as a cause of infections in humans (Andrade-Figueiredo and Leal-Balbino, 2016) and have been also isolated from animals (Phaku et al., 2016). Functional and genomic analysis of OS-MRSA and MRSA strains identified mutations in *femA*, a known auxiliary gene, as the possible cause of the observed decreased resistance to β-lactams (Giannouli et al., 2010; Phaku et al., 2016). Although being described many years after the emergence of pre-MRSA, the exact date of OS-MRSA emergence is uncertain. Actually, since for several decades detection of MRSA in many hospitals was based in purely phenotypic approaches, OS-MRSA may have passed unnoticed. It could be that OS-MRSA correspond to strains that have recently acquired *mecA* and that have developed alternative mechanisms to compensate for the cost of acquisition of an exogenous gene.

Altogether, data suggest that for acquiring and maintaining *mecA*, *S. aureus* strains had to adapt its genetic background, compensated for *mecA*/SCC*mec* fitness cost, or were already intrinsically equipped for it. Still it remains to be clarified

which genetic determinants and mechanism are involved in this adaptation process.

HOMOGENEOUS AND HETEROGENEOUS EXPRESSION OF RESISTANCE TO METHICILLIN

The genetic alterations in genetic backgrounds and associated metabolic alterations described above to have occurred upon SCC*mec* acquisition were frequently paralleled by alterations in the cell population profile of β-lactam resistance expression. Clinical MRSA isolates, when cultured, frequently exhibit a low level of methicillin resistance, but contain subpopulations of bacteria displaying very high levels of resistance to this antibiotic, a feature called heterogeneous resistance (Tomasz et al., 1991) (Step 10, **Figure 3**). Exposure of the hetero-MRSA strains to β-lactam antibiotics originates mutant strains in which all cells are uniformly highly resistant to β-lactams, named homogeneous methicillin resistance (Step 11, **Figure 3**) (Tomasz et al., 1991). Both hetero and homo resistance phenotypes can be found in clinical MRSA isolates, but they appear to correspond to two different and sequential evolutionary stages of β-lactams resistance expression. However, the molecular basis of the emergence of heterogeneous resistance and of the heterogeneous-to-homogeneous conversion is not totally understood and appears to derive by multiple different mechanism, of which only a few have yet been identified.

Genetic analysis of colonies within the highly resistant subpopulation of a heteregeneous MRSA strains, showed that high resistance was associated to the deletion of *lytH*, encoding a putative lytic enzyme homologous to a *N*-acetylmuramyl-L-alanine amidase (Fujimura and Murakami, 1997). But other mutations have been identified to provide the same type of phenotype, like mutations in *mecI* or in its promoter (Kondo et al., 2001). More recently, the comparison of the whole genome of strains selected from high and low level resistant subpopulations identified, in highly resistant strains, two additional mutations in *relA*, which is involved in the synthesis of (p)ppGppas, an effector of the stringent stress response to many environmental and genetic changes (Mwangi et al., 2013).

Ryffel et al. (1990) and Berger-Bachi and Rohrer (2002) first hypothesized that the heterogeneous-to-homogeneous conversion of methicillin resistance results from a spontaneous chromosomal mutation that is not linked to *mecA*. Kondo et al. (2001) showed by *in vitro trans*-complementation studies that *hmrA* and *hmrB*, which encode a putative aminohydrolase and an acyl carrier protein, respectively, were responsible for the conversion of the heterogenous profile (eagle type) of the N315 pre-MRSA strain into a uniformly highly resistant MRSA strain. Almost 20 years later a study wherein the whole genome of hetero-MRSA strain (N315) and its derivative homogeneously resistant strain selected by imipenem exposure,

were compared confirmed that they differed in a single non-synonymous mutation in *rpoB*, encoding the RNA polymerase β subunit (Aiba et al., 2013). Furthermore more recently, WGS revealed that tandem amplification of the SCC*mec* near its integration site was another alternative mechanism driving the heterogenous-to-homogeneous conversion (Gallagher et al., 2017).

CONCLUDING REMARKS

The development of *mecA*-mediated resistance to β-lactams was induced by human use of β-lactam antibiotics both to treat human infections and feed additives and involved several key genetic events: (1) the evolution of a native gene into a resistance determinant occurring at the native location; (2) the evolution of the SCC elements occurring at the *orfX* region; (3) integration of the *mec* complex and neighboring regions into a SCC element; (4) the adaptation of the host bacteria genetic background; (5) dissemination of SCC*mec* among staphylococci colonizing animals; (6) dissemination of SCC*mec* among staphylococci colonizing both animals and humans. Strikingly, most of the events that lead to β-lactam resistance development have occurred within the group of the most primitive animal-related *Staphylococcus* species isolated from production animals or human infection, suggesting it was a bacterial survival strategy against the human use of antimicrobials. The jump of SCC*mec* from animal to human-associated *Staphylococcus* species, like *S. aureus*, was a key event leading to several worldwide pandemics.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

REFERENCES

Aiba, Y., Katayama, Y., Hishinuma, T., Murakami-Kuroda, H., and Cui, L., Hiramatsu, K. (2013). Mutation of RNA polymerase beta-subunit gene promotes heterogeneous-to-homogeneous conversion of beta-lactam resistance in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 57, 4861–4871. doi: 10.1128/AAC.00720-13

Aminov, R. (2017). History of antimicrobial drug discovery: major classes and health impact. *Biochem. Pharmacol.* 133, 4–19. doi: 10.1016/j.bcp.2016.10.001

Andrade-Figueiredo, M., and Leal-Balbino, T. C. (2016). Clonal diversity and epidemiological characteristics of *Staphylococcus aureus*: high prevalence of oxacillin-susceptible *mecA*-positive *Staphylococcus aureus* (OS-MRSA) associated with clinical isolates in Brazil. *BMC Microbiol.* 16:115. doi: 10.1186/s12866-016-0733-4

Antignac, A., and Tomasz, A. (2009). Reconstruction of the phenotypes of methicillin-resistant *Staphylococcus aureus* by replacement of the staphylococcal cassette chromosome *mec* with a plasmid-borne copy of *Staphylococcus sciuri pbpD* gene. *Antimicrob. Agents Chemother.* 53, 435–441. doi: 10.1128/AAC.01099-08

Archer, G. L., and Niemeyer, D. M. (1994). Origin and evolution of DNA associated with resistance to methicillin in staphylococci. *Trends Microbiol.* 2, 343–347. doi: 10.1016/0966-842X(94)90608-4

Arede, P., Milheirico, C., de Lencastre, H., and Oliveira, D. C. (2012). The anti-repressor MecR2 promotes the proteolysis of the *mecA* repressor and enables optimal expression of beta-lactam resistance in MRSA. *PLoS Pathog.* 8:e1002816. doi: 10.1371/journal.ppat.1002816

Armand-Lefevre, L., Ruimy, R., and Andremont, A. (2005). Clonal comparison of *Staphylococcus aureus* isolates from healthy pig farmers, human controls, and pigs. *Emerg. Infect. Dis.* 11, 711–714. doi: 10.3201/eid1105.040866

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., and Hiramatsu, K. (2009). Complete genome sequence of *Macrococcus caseolyticus* strain JCSCS5402, [corrected] reflecting the ancestral genome of the human-pathogenic staphylococci. *J. Bacteriol.* 191, 1180–1190. doi: 10.1128/JB.01058-08

Baig, S., Johannesen, T. B., Overballe-Petersen, S., Larsen, J., Larsen, A. R., and Stegger, M. (2018). Novel SCC*mec* type XIII (9A) identified in an ST152 methicillin-resistant *Staphylococcus aureus*. *Infect. Genet. Evol.* 61, 74–76. doi: 10.1016/j.meegid.2018.03.013

Barber, M., and Rozwadowska-Dowzenko, M. (1948). Infection by penicillin-resistant staphylococci. *Lancet* 2, 641–644. doi: 10.1016/S0140-6736(48)92166-7

Becker, K., Ballhausen, B., Kock, R., and Kriegeskorte, A. (2014). Methicillin resistance in *Staphylococcus* isolates: the "*mec* alphabet" with specific consideration of *mecC*, a *mec* homolog associated with zoonotic *S. aureus* lineages. *Int. J. Med. Microbiol.* 304, 794–804. doi: 10.1016/j.ijmm.2014.06.007

Becker, K., van Alen, S., Idelevich, E. A., Schleimer, N., Seggewiss, J., Mellmann, A., et al. (2018). Plasmid-encoded transferable *mecB*-mediated methicillin resistance in *Staphylococcus aureus*. *Emerg. Infect. Dis.* 24, 242–248. doi: 10.3201/eid2402.171074

Berger-Bachi, B., and Rohrer, S. (2002). Factors influencing methicillin resistance in staphylococci. *Arch. Microbiol.* 178, 165–171. doi: 10.1007/s00203-002-0436-0

Berglund, C., Ito, T., Ikeda, M., Ma, X. X., Soderquist, B., and Hiramatsu, K. (2008). Novel type of staphylococcal cassette chromosome *mec* in a methicillin-resistant *Staphylococcus aureus* strain isolated in Sweden. *Antimicrob. Agents Chemother.* 52, 3512–3516. doi: 10.1128/AAC.00087-08

Bouchami, O., Ben Hassen, A., de Lencastre, H., and Miragaia, M. (2010). High prevalence of *mec* complex C and *ccrC* is independent of SCC*mec* type V in *Staphylococcus haemolyticus*. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 605–614. doi: 10.1007/s10096-011-1354-3

Bouchami, O., Ben Hassen, A., de Lencastre, H., and Miragaia, M. (2012). Molecular epidemiology of methicillin-resistant *Staphylococcus hominis* (MRSHo): low clonality and reservoirs of SCC*mec* structural elements. *PLoS One* 6:e21940. doi: 10.1371/journal.pone.0021940

Boundy, S., Safo, M. K., Wang, L., Musayev, F. N., O'Farrell, H. C., Rife, J. P., et al. (2013). Characterization of the *Staphylococcus aureus* rRNA methyltransferase encoded by *orfX*, the gene containing the staphylococcal chromosome Cassette *mec* (SCC*mec*) insertion site. *J. Biol. Chem.* 288, 132–140. doi: 10.1074/jbc.M112.385138

Calazans-Silva, A. C., Medeiros, P. T., Araujo, D. M., Carvalho, B. O., Coelho, I. S., Coelho, S. M., et al. (2014). Genetic analysis of *mecA* gene and detection of homologue *pbpD* in *Staphylococcus sciuri* group. *Braz. J. Microbiol.* 45, 651–655. doi: 10.1590/S1517-83822014000200038

Castanon, J. I. (2007). History of the use of antibiotic as growth promoters in European poultry feeds. *Poult. Sci.* 86, 2466–2471. doi: 10.3382/ps.2007-00249

Chambers, H. F., and Deleo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* 7, 629–641. doi: 10.1038/nrmicro2200

Chlebowicz, M. A., Maslanova, I., Kuntova, L., Grundmann, H., Pantucek, R., Doskar, J., et al. (2014). The Staphylococcal Cassette Chromosome *mec* type V from *Staphylococcus aureus* ST398 is packaged into bacteriophage capsids. *Int. J. Med. Microbiol.* 304, 764–774. doi: 10.1016/j.ijmm.2014.05.010

Chongtrakool, P., Ito, T., Ma, X. X., Kondo, Y., Trakulsomboon, S., Tiensasitorn, C., et al. (2006). Staphylococcal cassette chromosome *mec* (SCC*mec*) typing of methicillin-resistant *Staphylococcus aureus* strains isolated in 11 Asian countries: a proposal for a new nomenclature for SCC*mec* elements. *Antimicrob. Agents Chemother.* 50, 1001–1012. doi: 10.1128/AAC.50.3.1001-1012.2006

Cohen, S., and Sweeney, H. M. (1970). Transduction of methicillin resistance in *Staphylococcus aureus* dependent on an unusual specificity of the recipient strain. *J. Bacteriol.* 104, 1158–1167.

Continuous Discharge Certificate [CDC] (1999). Four pediatric deaths from community-acquired methicillin resistant *Staphylococcus aureus* - Minnesota and North Dakota, 1997-1999. *MMWR* 48, 707–710.

Couto, I., de Lencastre, H., Severina, E., Kloos, W., Webster, J. A., Hubner, R. J., et al. (1996). Ubiquitous presence of a *mecA* homologue in natural isolates of *Staphylococcus sciuri*. *Microb. Drug Resist.* 2, 377–391. doi: 10.1089/mdr.1996.2.377

Couto, I., Wu, S. W., Tomasz, A., and de Lencastre, H. (2003). Development of methicillin resistance in clinical isolates of *Staphylococcus sciuri* by transcriptional activation of the *mecA* homologue native to s. *J. Bacteriol.* 185, 645–653. doi: 10.1128/JB.185.2.645-653.2003

Crossley, K. B., and Archer, G. L. (1997). *The Staphylococci in Human Disease*. London: Churchill Livingstone.

de Lencastre, H., and Tomasz, A. (1994). Reassessment of the number of auxiliary genes essential for expression of high-level methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 38, 2590–2598. doi: 10.1128/AAC.38.11.2590

Deurenberg, R. H., and Stobberingh, E. E. (2008). The evolution of *Staphylococcus aureus*. *Infect. Genet. Evol.* 8, 747–763. doi: 10.1016/j.meegid.2008.07.007

Deurenberg, R. H., and Stobberingh, E. E. (2009). The molecular evolution of hospital- and community-associated methicillin-resistant *Staphylococcus aureus*. *Curr. Mol. Med.* 9, 100–115. doi: 10.2174/156652409787581637

Diep, B. A., Stone, G. G., Basuino, L., Graber, C. J., Miller, A., des Etages, S. A., et al. (2008). The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* 197, 1523–1530. doi: 10.1086/587907

Ender, M., McCallum, N., Adhikari, R., and Berger-Bachi, B. (2004). Fitness cost of SCC*mec* and methicillin resistance levels in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 48, 2295–2297. doi: 10.1128/AAC.48.6.2295-2297.2004

Entenza, J. M., Hohl, P., Heinze-Krauss, I., Glauser, M. P., and Moreillon P. (2002). BAL9141, a novel extended-spectrum cephalosporin active against methicillin-resistant *Staphylococcus aureus* in treatment of experimental endocarditis. *Antimicrob. Agents Chemother.* 46, 171–177. doi: 10.1128/AAC.46.1.171-177.2002

European Union [EU] (2003). REGULATION (EC) No 1831/2003 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on additives for use in animal nutrition. *Off. J. Eur. Union.* L 268/29, 1–30. (accessed Dec 30, 2005).

Figueiredo, T. A., Ludovice, A. M., and Sobral, R. G. (2014). Contribution of peptidoglycan amidation to beta-lactam and lysozyme resistance in different genetic lineages of *Staphylococcus aureus*. *Microb. Drug Resist.* 20, 238–249. doi: 10.1089/mdr.2014.0042

Fleming, A. (1929). Classics in infectious diseases: on the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation

of *B. influenzae* by Alexander Fleming, Reprinted from the British Journal of Experimental Pathology 10:226-236, 1929. *Rev. Infect. Dis.* 2, 129–139. doi: 10.1093/clinids/2.1.129

Fluit, A. C. (2012). Livestock-associated *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 18, 735–744. doi: 10.1111/j.1469-0691.2012.03846.x

Food and Drug Administration [FDA] (2018). *Antimicrobial Resistance.* Avilable at: http://www.fda.gov/animalveterinary/safetyhealth/antimicrobialresistance/default.htm

Fuda, C., Suvorov, M., Vakulenko, S. B., and Mobashery, S. (2004). The basis for resistance to beta-lactam antibiotics by penicillin-binding protein 2a of methicillin-resistant *Staphylococcus aureus*. *J. Biol. Chem.* 279, 40802–40806. doi: 10.1074/jbc.M403589200

Fujimura, T., and Murakami, K. (1997). Increase of methicillin resistance in *Staphylococcus aureus* caused by deletion of a gene whose product is homologous to lytic enzymes. *J. Bacteriol.* 179, 6294–6301. doi: 10.1128/jb.179.20.6294-6301.1997

Gallagher, L. A., Coughlan, S., Black, N. S., Lalor, P., Waters, E. M., Wee, B., et al. (2017). Tandem amplification of the staphylococcal cassette chromosome *mec* element can drive high-level methicillin resistance in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 61:e00869-17. doi: 10.1128/AAC.00869-17

Garcia-Alvarez, L., Holden, M. T., Lindsay, H., Webb, C. R., Brown, D. F., Curran, M. D., et al. (2011). Meticillin-resistant *Staphylococcus aureus* with a novel *mecA* homologue in human and bovine populations in the UK and Denmark: a descriptive study. *Lancet Infect. Dis.* 11, 595–603. doi: 10.1016/S1473-3099(11)70126-8

Ghuysen, J. M. (1991). Serine beta-lactamases and penicillin-binding proteins. *Annu. Rev. Microbiol.* 45, 37–67. doi: 10.1146/annurev.mi.45.100191.000345

Ghuysen, J. M. (1994). Molecular structures of penicillin-binding proteins and beta-lactamases. *Trends Microbiol.* 2, 372–380. doi: 10.1016/0966-842X(94)90614-9

Giannouli, S., Labrou, M., Kyritsis, A., Ikonomidis, A., Pournaras, S., Stathopoulos, C., et al. (2010). Detection of mutations in the FemXAB protein family in oxacillin-susceptible *mecA*-positive *Staphylococcus aureus* clinical isolates. *J. Antimicrob. Chemother.* 65, 626–633. doi: 10.1093/jac/dkq039

Gomez-Sanz, E., Schwendener, S., Thomann, A., Gobeli Brawand, S., and Perreten, V. (2015). First staphylococcal cassette chromosome *mec* containing a *mecB*-carrying gene complex independent of transposon Tn6045 in a *Macrococcus caseolyticus* isolate from a canine infection. *Antimicrob. Agents Chemother.* 59, 4577–4583. doi: 10.1128/AAC.05064-14

Grave, K., Greko, C., Kvaale, M. K., Torren-Edo, J., Mackay, D., Muller, A., et al. (2012). Sales of veterinary antibacterial agents in nine European countries during 2005-09: trends and patterns. *J. Antimicrob. Chemother.* 67, 3001–3008. doi: 10.1093/jac/dks298

Grundmann, H., Aires-de-Sousa, M., Boyce, J., and Tiemersma, E. (2006). Emergence and resurgence of meticillin-resistant *Staphylococcus aureus* as a public-health threat. *Lancet* 368, 874–885. doi: 10.1016/S0140-6736(06)68853-3

Haaber, J., Leisner, J. J., Cohn, M. T., Catalan-Moreno, A., Nielsen, J. B., Westh, H., et al. (2016). Bacterial viruses enable their host to acquire antibiotic resistance genes from neighbouring cells. *Nat. Commun.* 7:13333. doi: 10.1038/ncomms13333

Hao, H., Cheng, G., Iqbal, Z., Ai, X., Hussain, H. I., Huang, L., et al. (2014). Benefits and risks of antimicrobial use in food-producing animals. *Front. Microbiol.* 5:288. doi: 10.3389/fmicb.2014.00288

Harris, S. R., Feil, E. J., Holden, M. T., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2012). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395

Harrison, E. M., Paterson, G. K., Holden, M. T., Ba, X., Rolo, J., Morgan, F. J., et al. (2014). A novel hybrid SCC*mec-mecC* region in *Staphylococcus sciuri*. *J. Antimicrob. Chemother.* 69, 911–918. doi: 10.1093/jac/dkt452

Harrison, E. M., Paterson, G. K., Holden, M. T., Morgan, F. J., Larsen, A. R., Petersen, A., et al. (2013). A *Staphylococcus xylosus* isolate with a new *mecC* allotype. *Antimicrob. Agents Chemother.* 57, 1524–1528. doi: 10.1128/AAC.01882-12

Hartman, B. J., and Tomasz, A. (1984). Low-affinity penicillin-binding protein associated with beta-lactam resistance in *Staphylococcus aureus*. *J. Bacteriol.* 158, 513–516.

Hauschild, T., Stepanovic, S., and Zakrzewska-Czerwinska, J. (2010). *Staphylococcus stepanovicii* sp. nov., a novel novobiocin-resistant oxidase-positive staphylococcal species isolated from wild small mammals. *Syst. Appl. Microbiol.* 33, 183–187. doi: 10.1016/j.syapm.2010.03.004

Herold, B. C., Immergluck, L. C., Maranan, M. C., Lauderdale, D. S., Gaskin, R. E., Boyle-Vavra, S., et al. (1998). Community-acquired methicillin-resistant *Staphylococcus aureus* in children with no identified predisposing risk. *JAMA* 279, 593–598. doi: 10.1001/jama.279.8.593

Higgins, P. G., Rosato, A. E., Seifert, H., Archer, G. L., and Wisplinghoff, H. (2009). Differential expression of *ccrA* in methicillin-resistant *Staphylococcus aureus* strains carrying staphylococcal cassette chromosome *mec* type II and IVa elements. *Antimicrob. Agents Chemother.* 53, 4556–4558. doi: 10.1128/AAC.00395-09

Hiramatsu, K., Asada, K., Suzuki, E., Okonogi, K., and Yokota, T. (1992). Molecular cloning and nucleotide sequence determination of the regulator region of *mecA* gene in methicillin-resistant *Staphylococcus aureus* (MRSA). *FEBS Lett.* 298, 133–136. doi: 10.1016/0014-5793(92)80039-J

Hiramatsu, K., Ito, T., Tsubakishita, S., Sasaki, T., Takeuchi, F., Morimoto, Y., et al. (2013). Genomic basis for methicillin resistance in *Staphylococcus aureus*. *Infect. Chemother.* 45, 117–136. doi: 10.3947/ic.2013.45.2.117

Holden, M. T., Feil, E. J., Lindsay, J. A., Peacock, S. J., Day, N. P., Enright, M. C., et al. (2004). Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9786–9791. doi: 10.1073/pnas.0402521101

Huijbers, P. M., Blaak, H., de Jong, M. C., Graat, E. A., Vandenbroucke-Grauls, C. M., and de Roda Husman, A. M. (2015). Role of the environment in the transmission of antimicrobial resistance to humans: a review. *Environ. Sci. Technol.* 49, 11993–12004. doi: 10.1021/acs.est.5b02566

International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements [IWG-SCC], Ito, T. K., Hiramatsu, D. C., Oliveira, H., de Lencastre, K., Zhang, H., et al. (2009). Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. *Antimicrob. Agents Chemother.* 53, 4961–4967. doi: 10.1128/AAC.00579-09

Ishikawa, T., Matsunaga, N., Tawada, H., Kuroda, N., Nakayama, Y., Ishibashi, Y., et al. (2003). TAK-599, a novel N-phosphono type prodrug of anti-MRSA cephalosporin T-91825: synthesis, physicochemical and pharmacological properties. *Bioorg. Med. Chem.* 11, 2427–2437. doi: 10.1016/S0968-0896(03)00126-3

Ito, T., Katayama, Y., Asada, K., Mori, N., Tsutsumimoto, K., Tiensasitorn, C., et al. (2001). Structural comparison of three types of staphylococcal cassette chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 45, 1323–1336. doi: 10.1128/AAC.45.5.1323-1336.2001

Ito, T., Ma, X. X., Takeuchi, F., Okuma, K., Yuzawa, H., and Hiramatsu, K. (2004). Novel type V staphylococcal cassette chromosome *mec* driven by a novel cassette chromosome recombinase, *ccrC*. *Antimicrob. Agents Chemother.* 48, 2637–2651. doi: 10.1128/AAC.48.7.2637-2651.2004

Jevons, M. P. (1961). Celbenin-resistant' staphylococci. *Br. Med. J.* 1, 124–125. doi: 10.1136/bmj.1.5219.124-a

Katayama, Y., Ito, T., and Hiramatsu, K. (2000). A new class of genetic element, *Staphylococcus* cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 44, 1549–1555. doi: 10.1128/AAC.44.6.1549-1555.2000

Katayama, Y., Ito, T., and Hiramatsu, K. (2001). Genetic organization of the chromosome region surrounding *mecA* in clinical staphylococcal strains: role of IS*431*-Mediated *mecI* deletion in expression of resistance in *mecA*-carrying, low-level methicillin- resistant *Staphylococcus haemolyticus*. *Antimicrob. Agents Chemother.* 45, 1955–1963. doi: 10.1128/AAC.45.7.1955-1963.2001

Katayama, Y., Takeuchi, F., Ito, T., Ma, X. X., Ui-Mizutani, Y., Kobayashi, I., et al. (2003a). Identification in methicillin-susceptible *Staphylococcus hominis* of an active primordial mobile genetic element for the staphylococcal cassette

chromosome *mec* of methicillin-resistant *Staphylococcus aureus. J. Bacteriol.* 185, 2711–2722. doi: 10.1128/JB.185.9.2711-2722.2003

Katayama, Y., Zhang, H. Z., Hong, D., and Chambers, H. F. (2003b). Jumping the barrier to beta-lactam resistance in *Staphylococcus aureus. J. Bacteriol.* 185, 5465–5472. doi: 10.1128/JB.185.18.5465-5472.2003

Kaya, H., Hasman, H., Larsen, J., Stegger, M., Johannesen, T. B., Allesoe, R. L., et al. (2018). SCC*mec*Finder, a Web-Based tool for typing of staphylococcal cassette chromosome *mec* in *Staphylococcus aureus* using whole-genome sequence data. *mSphere* 3:e00612-17.

Kim, C., Milheirico, C., Gardete, S., Holmes, M. A., Holden, M. T., de Lencastre, H., et al. (2012). Properties of a novel PBP2A protein homolog from *Staphylococcus aureus* strain LGA251 and its contribution to the beta-lactam-resistant phenotype. *J. Biol. Chem.* 287, 36854–36863. doi: 10.1074/jbc.M112.395962

Kloos, W. (1986). "Ecology of human skin," in *Coagulase-Negative Staphylococci*, eds P.-A. Mardh and K. H. Schleifer (Stockholm: Almqvist & Wiksell).

Kloos, W. (1997). "Taxonomy and systematics of staphylococci indigenous to humans," in *The Staphylococci in Human Disease*, ed. K. B. Crossley (New York, NY: Churchill Livingstone), 113–137.

Kloos, W. E. (1980). Natural populations of the genus *Staphylococcus. Annu. Rev. Microbiol.* 34, 559–592. doi: 10.1146/annurev.mi.34.100180.003015

Kondo, N., Kuwahara-Arai, K., Kuroda-Murakami, H., Tateda-Suzuki, E., and Hiramatsu, K. (2001). Eagle-type methicillin resistance: new phenotype of high methicillin resistance under *mec* regulator gene control. *Antimicrob. Agents Chemother.* 45, 815–824. doi: 10.1128/AAC.45.3.815-824.2001

Kuwahara-Arai, K., Kondo, N., Hori, S., Tateda-Suzuki, E., and Hiramatsu, K. (1996). Suppression of methicillin resistance in a *mecA*-containing pre-methicillin-resistant *Staphylococcus aureus* strain is caused by the *mecI*-mediated repression of PBP 2' production. *Antimicrob. Agents Chemother.* 40, 2680–2685. doi: 10.1128/AAC.40.12.2680

Li, S., Skov, R. L., Han, X., Larsen, A. R., Larsen, J., Sorum, M., et al. (2011). Novel types of staphylococcal cassette chromosome *mec* elements identified in clonal complex 398 methicillin-resistant *Staphylococcus aureus* strains. *Antimicrob. Agents Chemother.* 55, 3046–3050. doi: 10.1128/AAC.01475-10

Lim, D., and Strynadka, N. C. (2002). Structural basis for the beta lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus. Nat. Struct. Biol.* 9, 870–876.

Lindsay, J. A. (2013). Hospital-associated MRSA and antibiotic resistance-what have we learned from genomics? *Int. J. Med. Microbiol.* 303, 318–323. doi: 10.1016/j.ijmm.2013.02.005

Loncaric, I., Kubber-Heiss, A., Posautz, A., Stalder, G. L., Hoffmann, D., Rosengarten, R., et al. (2013). Characterization of methicillin-resistant *Staphylococcus* spp. carrying the *mecC* gene, isolated from wildlife. *J. Antimicrob. Chemother.* 68, 2222–2225. doi: 10.1093/jac/dkt186

Luong, T. T., Ouyang, S., Bush, K., and Lee, C. Y. (2002). Type 1 capsule genes of *Staphylococcus aureus* are carried in a staphylococcal cassette chromosome genetic element. *J. Bacteriol.* 184, 3623–3629. doi: 10.1128/JB.184.13.3623-3629.2002

Ma, X. X., Ito, T., Tiensasitorn, C., Jamklang, M., Chongtrakool, P., Boyle-Vavra, S., et al. (2002). Novel type of staphylococcal cassette chromosome *mec* identified in community-acquired methicillin-resistant *Staphylococcus aureus* strains. *Antimicrob. Agents Chemother.* 46, 1147–1152. doi: 10.1128/AAC.46.4.1147-1152.2002

Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O., and Dessen, A. (2006). Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes. *FEMS Microbiol. Rev.* 30, 673–691. doi: 10.1111/j.1574-6976.2006.00024.x

Maslanova, I., Doskar, J., Varga, M., Kuntova, L., Muzik, J., Maluskova, D., et al. (2013). Bacteriophages of *Staphylococcus aureus* efficiently package various bacterial genes and mobile genetic elements including SCC*mec* with different frequencies. *Environ. Microbiol. Rep.* 5, 66–73. doi: 10.1111/j.1758-2229.2012.00378.x

Memmi, G., Filipe, S. R., Pinho, M. G., Fu, Z., and Cheung, A. (2008). *Staphylococcus aureus* PBP4 is essential for beta-lactam resistance in community-acquired methicillin-resistant strains. *Antimicrob. Agents Chemother.* 52, 3955–3966. doi: 10.1128/AAC.00049-08

Miragaia, M., Thomas, J. C., Couto, I., Enright, M. C., and de Lencastre, H. (2007). Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing (MLST) data. *J. Bacteriol.* 189, 2540–2552. doi: 10.1128/JB.01484-06

Mir-Sanchis, I., Roman, C. A., Misiura, A., Pigli, Y. Z., Boyle-Vavra, S., and Rice, P. A. (2016). Staphylococcal SCC*mec* elements encode an active MCM-like helicase and thus may be replicative. *Nat. Struct. Mol. Biol.* 23, 891–898. doi: 10.1038/nsmb.3286

Monecke, S., Muller, E., Schwarz, S., Hotzel, H., and Ehricht, R. (2012). Rapid microarray-based identification of different *mecA* alleles in Staphylococci. *Antimicrob. Agents Chemother.* 56, 5547–5554. doi: 10.1128/AAC.00574-12

Mongkolrattanothai, K., Boyle, S., Murphy, T. V., and Daum, R. S. (2004). Novel non-*mecA*-containing staphylococcal chromosomal cassette composite island containing *pbp4* and *tagF* genes in a commensal staphylococcal species: a possible reservoir for antibiotic resistance islands in *Staphylococcus aureus. Antimicrob. Agents Chemother.* 48, 1823–1836. doi: 10.1128/AAC.48.5.1823-1836.2004

Moran, G. J., Krishnadasan, A., Gorwitz, R. J., Fosheim, G. E., McDougal, L. K., Carey, R. B., et al. (2006). Methicillin-resistant *S. aureus* infections among patients in the emergency department. *N. Engl. J. Med.* 355, 666–674. doi: 10.1056/NEJMoa055356

Morikawa, K., Takemura, A. J., Inose, Y., Tsai, M., Nguyen Thi le, T., Ohta, T., et al. (2012). Expression of a cryptic secondary sigma factor gene unveils natural competence for DNA transformation in *Staphylococcus aureus. PLoS Pathog.* 8:e1003003. doi: 10.1371/journal.ppat.1003003

Mwangi, M. M., Kim, C., Chung, M., Tsai, J., Vijayadamodar, G., Benitez, M., et al. (2013). Whole-genome sequencing reveals a link between beta-lactam resistance and synthetases of the alarmone (p)ppGpp in *Staphylococcus aureus. Microb. Drug Resist.* 19, 153–159. doi: 10.1089/mdr.2013.0053

National Research Council (1980). *Committee to Study the Human Health Effects of Subtherapeutic Antibiotic Use in Animal Feeds. Antibiotics In Animal Feeds, National Research Council (US) Committee to Study the Human Health Effects of Subtherapeutic Antibiotic Use in Animal Feeds The Effects on Human Health of Subtherapeutic Use of Antimicrobials in Animal Feeds*. Washington, DC: National Academies Press.

Navarre, W. W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S. J., et al. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella. Science* 313, 236–238. doi: 10.1126/science.1128794

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500

Oliveira, D. C., Milheirico, C., and de Lencastre, H. (2006). Redefining a structural variant of staphylococcal cassette chromosome *mec*, SCC*mec* type VI. *Antimicrob. Agents Chemother.* 50, 3457–3459. doi: 10.1128/AAC.00629-06

Oliveira, D. C., Tomasz, A., and de Lencastre, H. (2001). The evolution of pandemic clones of methicillin-resistant *Staphylococcus aureus*: identification of two ancestral genetic backgrounds and the associated mec elements. *Microb. Drug Resist.* 7, 349–361. doi: 10.1089/10766290152773365

Oliveira, D. C., Tomasz, A., and de Lencastre, H. (2002). Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant *Staphylococcus aureus. Lancet Infect. Dis.* 2, 180–189. doi: 10.1016/S1473-3099(02)00227-X

Otto, M. (2013). Community-associated MRSA: what makes them special? *Int. J. Med. Microbiol.* 303, 324–330. doi: 10.1016/j.ijmm.2013.02.007

Phaku, P., Lebughe, M., Strauss, L., Peters, G., Herrmann, M., Mumba, D., et al. (2016). Unveiling the molecular basis of antimicrobial resistance in *Staphylococcus aureus* from the Democratic Republic of the Congo using whole genome sequencing. *Clin. Microbiol. Infect.* 22:644 e1–644.e5.

Pinho, M. G., de Lencastre, H., and Tomasz, A. (2001). An acquired and a native penicillin-binding protein cooperate in building the cell wall of drug-resistant staphylococci. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10886–10891. doi: 10.1073/pnas.191260798

Rao, Q., Shang, W., Hu, X., and Rao, X. (2015). *Staphylococcus aureus* ST121: a globally disseminated hypervirulent clone. *J. Med. Microbiol.* 64, 1462–1473. doi: 10.1099/jmm.0.000185

Ray, M. D., Boundy, S., and Archer, G. L. (2016). Transfer of the methicillin resistance genomic island among staphylococci by conjugation. *Mol. Microbiol.* 100, 675–685. doi: 10.1111/mmi.13340

report SE (2010/2015). Sales of veterinary antimicrobial agents in 30 European countries in 2015*.

Robinson, D. A., and Enright, M. C. (2003). Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 47, 3926–3934. doi: 10.1128/AAC.47.12.3926-3934. 2003

Rolo, J., de Lencastre, H., and Miragaia, M. (2012). Strategies of adaptation of *Staphylococcus epidermidis* to hospital and community: amplification and diversification of SCC*mec*. *J. Antimicrob. Chemother.* 67, 1333–1341. doi: 10. 1093/jac/dks068

Rolo, J., de Lencastre, H., and Miragaia, M. (2014). High frequency and diversity of cassette chromosome recombinases (ccr) in methicillin-susceptible *Staphylococcus sciuri*. *J. Antimicrob. Chemother.* 69, 1461–1469. doi: 10.1093/ jac/dku028

Rolo, J., Worning, P., Nielsen, J. B., Bowden, R., Bouchami, O., Damborg, P., et al. (2017a). Evolutionary origin of the staphylococcal cassette chromosome *mec* (SCC*mec*). *Antimicrob. Agents Chemother.* 61:e02302-16. doi: 10.1128/AAC. 02302-16

Rolo, J., Worning, P., Nielsen, J. B., Sobral, R., Bowden, R., Bouchami, O., et al. (2017b). Evidence for the evolutionary steps leading to *mecA*-mediated beta-lactam resistance in staphylococci. *PLoS Genet.* 13:e1006674. doi: 10.1371/ journal.pgen.1006674

Rountree, P. M., and Freeman, B. M. (1955). Infections caused by a particular phage type of *Staphylococcus aureus*. *Med. J. Aust.* 42, 157–161.

Ryffel, C., Tesch, W., Birch-Machin, I., Reynolds, P. E., Barberis-Maino, L., Kayser, F. H., et al. (1990). Sequence comparison of *mecA* genes isolated from methicillin-resistant *Staphylococcus aureus* and *Staphylococcus epidermidis*. *Gene* 94, 137–138. doi: 10.1016/0378-1119(90) 90481-6

Scharn, C. R., Tenover, F. C., and Goering, R. V. (2013). Transduction of staphylococcal cassette chromosome *mec* elements between strains of *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 57, 5233–5238. doi: 10. 1128/AAC.01058-13

Schleifer, K. H., Geyer, U., Kilpper-Balz, R., and Devriese, L. A. (1983). Elevation of *Staphylococcus sciuri* subsp. *lentus* (Kloos et al.) to species status: *Staphylococcus lentus* (Kloos et al.) comb. nov. *Syst. Appl. Microbiol.* 4, 382–387. doi: 10.1016/ S0723-2020(83)80022-8

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Sci. Rep.* 7:43797. doi: 10.1038/srep43797

Schwendener, S., and Perreten, V. (2018). The integrase of the *Macrococcus caseolyticus* resistance island *mecD* (McRI$_{mecD}$) inserts DNA site-specifically into *Staphylococcus* and *Bacillus* chromosomes. *Mol. Microbiol.* doi: 10.1111/ mmi.14112

Semmler, T., Harrison, E. M., Lubke-Becker, A., Ulrich, R. G., Wieler, L. H., Guenther, S., et al. (2016). A look into the melting pot: the *mecC*-harboring region is a recombination hot spot in *Staphylococcus stepanovicii*. *PLoS One* 11:e0147150. doi: 10.1371/journal.pone.0147150

Shahid, M., Sobia, F., Singh, A., Malik, A., Khan, H. M., Jonas, D., et al. (2009). Beta-lactams and beta-lactamase-inhibitors in current- or potential-clinical practice: a comprehensive update. *Crit. Rev. Microbiol.* 35, 81–108. doi: 10.1080/ 10408410902733979

Shore, A. C., Deasy, E. C., Slickers, P., Brennan, G., O'Connell, B., Monecke, S., et al. (2011a). Detection of staphylococcal cassette chromosome *mec* type XI carrying highly divergent *mecA, mecI, mecR1, blaZ*, and *ccr* genes in human clinical isolates of clonal complex 130 methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 55, 3765–3773. doi: 10.1128/AAC. 00187-11

Shore, A. C., Rossney, A. S., Brennan, O. M., Kinnevey, P. M., Humphreys, H., Sullivan, D. J., et al. (2011b). Characterization of a novel arginine catabolic mobile element (ACME) and staphylococcal chromosomal cassette *mec* composite island with significant homology to *Staphylococcus epidermidis* ACME type II in methicillin-resistant *Staphylococcus aureus* genotype ST22-MRSA-IV. *Antimicrob. Agents Chemother.* 55, 1896–1905. doi: 10.1128/AAC. 01756-10

Song, M. D., Wachi, M., Doi, M., Ishino, F., and Matsuhashi, M. (1987). Evolution of an inducible penicillin-target protein in methicillin-resistant *Staphylococcus aureus* by gene fusion. *FEBS Lett.* 221, 167–171. doi: 10.1016/0014-5793(87) 80373-3

Stojanov, M., Sakwinska, O., and Moreillon, P. (2013). Expression of SCCmec cassette chromosome recombinases in methicillin-resistant *Staphylococcus aureus* and *Staphylococcus epidermidis*. *J. Antimicrob. Chemother.* 68, 749–757. doi: 10.1093/jac/dks494

Suzuki, E., Kuwahara-Arai, K., Richardson, J. F., and Hiramatsu, K. (1993). Distribution of *mec* regulator genes in methicillin-resistant Staphylococcus clinical strains. *Antimicrob. Agents Chemother.* 37, 1219–1226. doi: 10.1128/ AAC.37.6.1219

Tomasz, A., Nachman, S., and Leaf, H. (1991). Stable classes of phenotypic expression in methicillin-resistant clinical isolates of staphylococci. *Antimicrob. Agents Chemother.* 35, 124–129. doi: 10.1128/AAC.35.1.124

Tsubakishita, S., Kuwahara-Arai, K., Baba, T., and Hiramatsu, K. (2010a). Staphylococcal cassette chromosome *mec*-like element in *Macrococcus caseolyticus*. *Antimicrob. Agents Chemother.* 54, 1469–1475. doi: 10.1128/AAC. 00575-09

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., and Hiramatsu, K. (2010b). Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. *Antimicrob. Agents Chemother.* 54, 4352–4359. doi: 10.1128/ AAC.00356-10

Vandenesch, F., Naimi, T., Enright, M. C., Lina, G., Nimmo, G. R., Heffernan, H., et al. (2003). Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. *Emerg. Infect. Dis.* 9, 978–984. doi: 10.3201/eid0908.030089

Vlack, S., Cox, L., Peleg, A. Y., Canuto, C., Stewart, C., Conlon, A., et al. (2006). Carriage of methicillin-resistant *Staphylococcus aureus* in a Queensland Indigenous community. *Med. J. Aust.* 184, 556–559.

Voss, A., Loeffen, F., Bakker, J., Klaassen, C., and Wulf, M. (2005). Methicillin-resistant *Staphylococcus aureus* in pig farming. *Emerg. Infect. Dis.* 11, 1965–1966. doi: 10.3201/eid1112.050428

Wang, L., Safo, M., and Archer, G. L. (2012). Characterization of DNA sequences required for the CcrAB-mediated integration of staphylococcal cassette chromosome *mec*, a *Staphylococcus aureus* genomic island. *J. Bacteriol.* 194, 486–498. doi: 10.1128/JB.05047-11

Waxman, D. J., and Strominger, J. L. (1983). Penicillin-binding proteins and the mechanism of action of beta-lactam antibiotics. *Annu. Rev. Biochem.* 52, 825–869. doi: 10.1146/annurev.bi.52.070183.004141

Westh, H., Zinn, C. S., and Rosdahl, V. T. (2004). An international multicenter study of antimicrobial consumption and resistance in *Staphylococcus aureus* isolates from 15 hospitals in 14 countries. *Microb. Drug Resist.* 10, 169–176. doi: 10.1089/1076629041310019

Wisplinghoff, H., Rosato, A. E., Enright, M. C., Noto, M., Craig, W., and Archer, G. L. (2003). Related clones containing SCC*mec* type IV predominate among clinically significant *Staphylococcus epidermidis* isolates. *Antimicrob. Agents Chemother.* 47, 3574–3579. doi: 10.1128/AAC.47.11.3574-3579. 2003

World Organization for Animal Health (2016). OIE Annual report on the use of antimicrobial agents in animals. Better Understanding of The Global Situation. OIE Annual report.

Wu, S., de Lencastre, H., and Tomasz, A. (1998). Genetic organization of the *mecA* region in methicillin-susceptible and methicillin-resistant strains of *Staphylococcus sciuri*. *J. Bacteriol.* 180, 236–242.

Wu, S., Piscitelli, C., de Lencastre, H., Tomasz, A. (1996). Tracking the evolutionary origin of the methicillin resistance gene: cloning and sequencing of a homologue of *mecA* from a methicillin susceptible strain of *Staphylococcus sciuri*. *Microb. Drug Resist.* 2, 435–441. doi: 10.1089/mdr.1996. 2.435

Wu, S. W., de Lencastre, H., and Tomasz, A. (2001). Recruitment of the *mecA* gene homologue of *Staphylococcus sciuri* into a resistance determinant and expression of the resistant phenotype in *Staphylococcus aureus*. *J. Bacteriol.* 183, 2417–2424. doi: 10.1128/JB.183.8.2417-2424.2001

Wu, Z., Li, F., Liu, D., Xue, H., and Zhao, X. (2015). Novel type XII Staphylococcal cassette chromosome *mec* harboring a new cassette chromosome recombinase, CcrC2. *Antimicrob. Agents Chemother.* 59, 7597–7601. doi: 10.1128/AAC. 01692-15

Zhang, K., McClure, J. A., Elsayed, S., Conly, J. M. (2009). Novel staphylococcal cassette chromosome *mec* type, tentatively designated type VIII, harboring class A *mec* and type 4 *ccr* gene complexes in a Canadian epidemic strain of methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 53, 531–540. doi: 10.1128/AAC.01118-08

Zhou, Y., Antignac, A., Wu, S. W., and Tomasz, A. (2008). Penicillin-binding proteins and cell wall composition in beta-lactam-sensitive and -resistant strains of *Staphylococcus sciuri*. *J. Bacteriol.* 190, 508–514. doi: 10.1128/JB.01549-07

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Laboratory-Based and Point-of-Care Testing for MSSA/MRSA Detection in the Age of Whole Genome Sequencing

Alex van Belkum[1]* and Olivier Rochas[2]

[1] Data Analytics Unit, bioMérieux, La Balme-les-Grottes, France, [2] Strategic Intelligence, Business Development Direction, bioMérieux, Marcy-l'Étoile, France

Staphylococcus aureus is an opportunistic pathogen of animals and humans that is capable of both colonizing and infecting its eukaryotic host. It is frequently detected in the clinical microbiology routine laboratory. *S. aureus* is capable of acquiring antibiotic resistance traits with ease and, given its rapid global dissemination, resistance to meticillin in *S. aureus* has received extensive coverage in the popular and medical press. The detection of meticillin-resistant versus meticillin-susceptible *S. aureus* (MRSA and MSSA) is of significant clinical importance. Detection of meticillin resistance is relatively straightforward since it is defined by a single determinant, penicillin-binding protein 2a', which exists in a limited number of genetic variants carried on various Staphylococcal Cassette Chromosomes *mec*. Diagnosis of MRSA and MSSA has evolved significantly over the past decades and there has been a strong shift from culture-based, phenotypic methods toward molecular detection, especially given the close correlation between the presence of the *mec* genes and phenotypic resistance. This brief review summarizes the current state of affairs concerning the mostly polymerase chain reaction-mediated detection of MRSA and MSSA in either the classical laboratory setting or at the point of care. The potential diagnostic impact of the currently emerging whole genome sequencing (WGS) technology will be discussed against a background of diagnostic, surveillance, and infection control parameters. Adequate detection of MSSA and MRSA is at the basis of any subsequent, more generic antibiotic susceptibility testing, epidemiological characterization, and detection of virulence factors, whether performed with classical technology or WGS analyses.

Keywords: *Staphylococcus aureus*, MSSA, MRSA, molecular testing, point of care, next generation sequencing, whole genome sequences

INTRODUCTION

Detection of infectious agents and their diseases is performed through a wide array of diagnostic methodologies. These range from *in silico* methods assessing host' susceptibility to colonization and infection (Suh et al., 2018) to direct or indirect detection of the pathogen itself. The latter tests are collectively known as *in vitro* diagnostics (IVD) and their execution requires a qualified laboratory environment and highly educated technicians and (clinical) microbiologists.

Next to the laboratory-based tests there are also more simple formats that should be safe to use outside of the laboratory by trained non-professionals at the point-of-need (PoN) or point-of-care (PoC). The more popular diagnostic tools are increasingly molecular in nature, having speed, specificity, and sensitivity superior to those of more classical, culture-based technologies. Molecular tests are based on different principles of which direct hybridization is among the most ancient. In addition, several different nucleic acid amplification technologies (NAATs) have been implemented [see Muldrew (2009) for a review]. Post-amplification processing often includes DNA fragment analysis and/or sequencing. Such tests are mostly aiming at detection and identification of disease-invoking bacterial species. Of note, primary detection and identification of micro-organisms are obvious pre-requisites to their further epidemiological characterization or research into their resistance and virulence characteristics. Complete diagnostic data sets can then be used for curing patients or for prevention of cross infection and infection control.

In the current era of multi-to-pan antibiotic resistance, there is an increasing interest in microbial tests that detect antibiotic resistance, one of the current medical scourges (Okeke et al., 2011; Kelly et al., 2016). Although phenotypic analysis is our heritage, optimal molecular diagnosis should allow for the simultaneous detection, identification, and genetic antibiotic susceptibility testing (AST) of infectious agents. DNA sequences

at the species level and at the level of resistance genes can be amplified at the same time in the same assay using the same clinical material as source (see **Figure 1** for a conceptual explanation). Indirect AST results in the detection of resistance markers and should lead to targeted treatment on the basis of the presumed activity of the product for which only the gene was detected. When innovative analytical techniques such as mass spectrometry, liquid chromatography, and nucleic acid sequencing are included advices on treatment may become more encompassing and cover-all.

Here, we will focus on the detection of meticillin resistance in the bacterial species *Staphylococcus aureus* as a model for the recent evolution of and the huge value of AST in clinical care. Meticillin-resistant *S. aureus* (MRSA) generates twice as much mortality than methicillin-susceptible *S. aureus* (MSSA) (Turnidge et al., 2009) and rapid molecular diagnostics has already been shown to reduce hospital stay and costs associated with MRSA infection (Brown and Paladino, 2010). Detecting both MSSA and MRSA is important since it guides therapeutic interventions with optimal antibiotics (Liu et al., 2011). Successful molecular diagnostic tests were developed by individual researchers (in house technology) but also by the IVD industry. In addition, quantification of the absolute number of bacterial cells in a clinical specimen is also important since different bacterial titers may be involved in colonization versus infection of human individuals. We

FIGURE 1 | Comparison between direct AST and indirect AST, showing the difference between the laboratory based and PoC protocols. Whereas the laboratory exercise generates precise minimum inhibitory concentrations (MICs), the indirect approach generates markers that need to be associated with therapeutic modalities through bio-informatics.
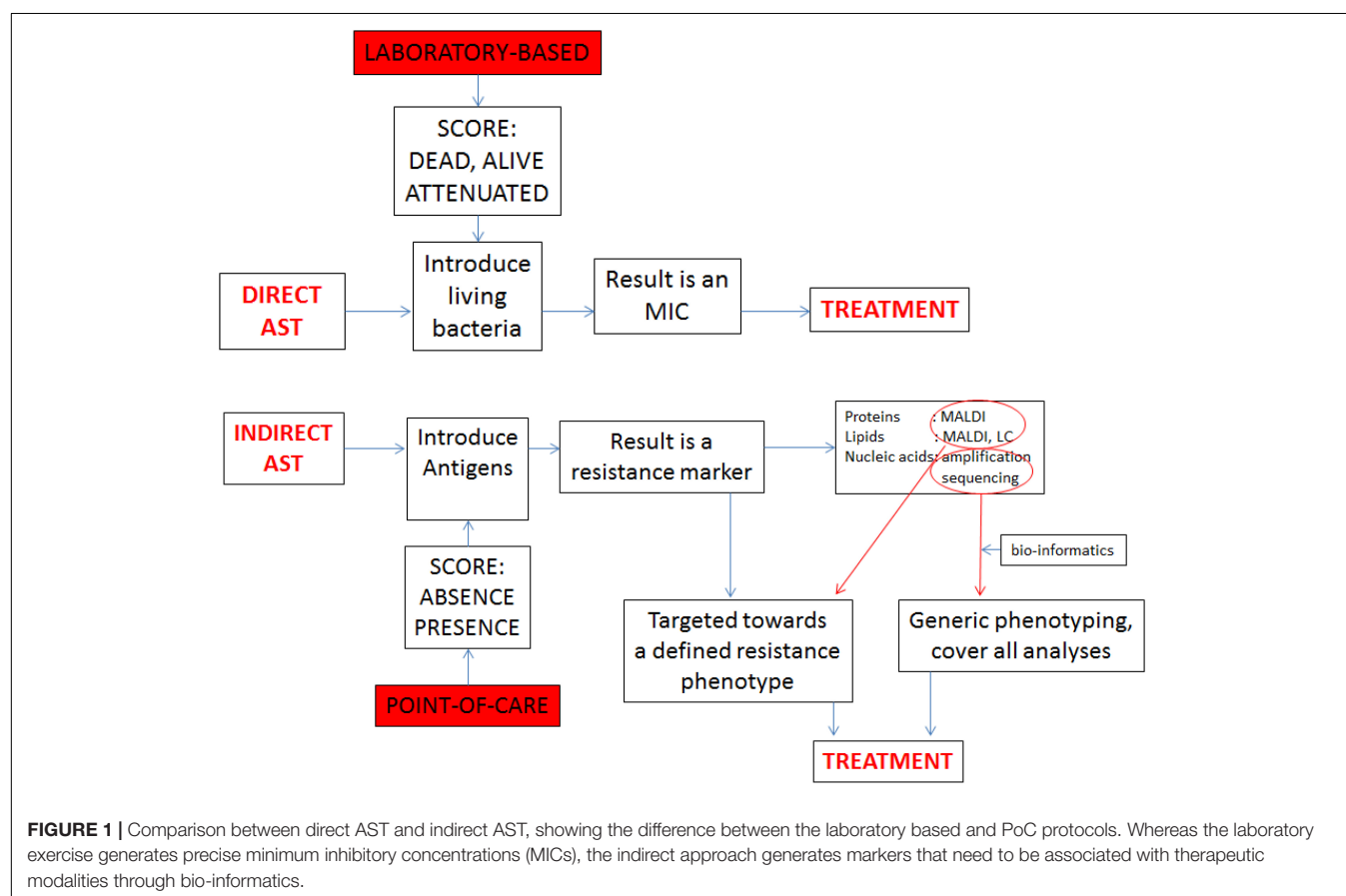
review nucleic acid-based tests that have been made recently available for the detection of MSSA and MRSA. When such tests are correctly used they do facilitate subsequent studies into the epidemiology, evolution, and spread of both MSSA and MRSA. Detailed assessment of resistance to a wider spectrum of antimicrobial agents can be performed and implementation of enhanced infection control becomes an option.

CULTURE-BASED DETECTION OF MSSA AND MRSA

There is no way of discussing molecular detection of MSSA and MRSA without briefly sketching the pre- and peri-molecular diagnostic landscape. Traditionally, staphylococcal colonization, and infection were diagnosed using culture-based technologies. These either employ generic, highly fertile culture media coupled to downstream bacterial species identification or species-specific enrichment media containing *S. aureus* selective components such as elevated salt concentrations. Addition of chromogenic compounds in the medium helps to identify *S. aureus* on the basis of colony morphology and color (Perry, 2017). Further taxonomic classification and identification of *S. aureus* can be done via agglutination assays or (commercially available) biochemical reactivity (e.g., API strips). Using combinations of simple phenotypic tests has been shown to allow for the adequate distinction of MSSA and MRSA (Verroken et al., 2016; Lüthje et al., 2017; Rees and Barr, 2017; Ábrók et al., 2018). Recent immunochromatographic methods such as the BinaxNOW (Alere, Scarborough, ME, United States) and the Clearview Exact PBP2a assay (Alere, Scarborough, ME, United States) have acquired a good position in the clinical laboratory given their modest price, rapidity, and good sensitivity and specificity (e.g., Kong et al., 2014). Still, modern laboratories now consider matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI ToF MS) as Gold Standard for staphylococcal identification (Bernardo et al., 2002; Szabados et al., 2010; Zhu et al., 2015). Distinguishing MRSA from MSSA using MALDI ToF MS is controversial with positive reports (Ueda et al., 2015; Rhoads et al., 2016; Sogawa et al., 2017) alternating with more negative ones (Du et al., 2002; Goldstein et al., 2013). Recent papers document the successful distinction between MSSA and MRSA, even including characterization of different MRSA clones (Østergaard et al., 2015; Zhang et al., 2015; Camoez et al., 2016). The main issue with these studies is that MALDI TOF MS detects proteins and that it is claimed that the *mec*A protein is produced in small amounts, often impossible to detect even by targeted mass spectrometry methods. So there is a significant risk that MALDI TOF MS will detect surrogate markers that also distinguish MRSA from MSSA. These may be markers of clonality rather than methicillin resistance and for this reason the collection of (preferably epidemiologically non-related) strains is of key importance. However, the overall impression is that MALDI ToF MS can be useful in the field of bacterial epidemiology but it will certainly not provide a universal "typing" solution.

Advanced MS methods, including for instance electron spray ionization (ESI) MS, may bring more universal solutions but these methods are too cumbersome, too time-consuming and too expensive at this stage (Charretier et al., 2015). Although MS essentially provides molecular testing, albeit at the protein level, NAATs still provide the best tool for distinguishing MSSA and MRSA.

SHORT INTRODUCTION INTO MOLECULAR TECHNOLOGIES

Using polymerase chain reaction (PCR), nucleic acid sequence-based amplification (NASBA), recombinase polymerase amplification (RPA), loop-mediated isothermal amplification (LAMP) and other systems, minute amounts of DNA can be amplified and detected using a variety of technological formats (Almassian et al., 2013; Yan et al., 2014). All of these methods have been shown to be useful in the detection of infectious agents and both PCR and several non-PCR tests have been commercialized (e.g., the LAMP-based Eazyplex test, see Henares et al., 2017). New methods surface regularly (e.g., those methods employing the specificity, sensitivity of enzymes involved in CRISPCas-mediated bacterial immunity to bacteriophages, see Gootenberg et al., 2017, 2018, for more details), several of them allowing genetic AST, and this import of new testing formats including their automatization will continue in the years to come.

There is a set of technologies that will undoubtedly have a huge impact on microbial detection and characterization: next generation sequencing (NGS) leading to the elucidation of the primary structure of complete bacterial chromosome sequences. Multiple elegant whole genome sequencing (WGS) NGS technologies have been developed three of which are currently commercialized and well-accessible to the diagnostic laboratory. Companies such as Illumina (San Diego, CA, United States), PacBio (Menlo Park, CA, United States), and Oxford Nanopore (Oxford, United Kingdom) provide exemplary methodologies suited for WGS. Further technical and usage detail on these methods will not be provided here but can be easily accessed through various recent reviews (Miyamoto et al., 2014; Quainoo et al., 2017; Rossen et al., 2017). NGS will find its way into the clinical microbiology routine laboratory over the years to come where it will fill in important niches in the rapid detection of pathogens and their epidemiological, antibiotic resistance, and virulence characteristics, possibly directly from clinical specimens. NGS will allow parallel sequencing of host DNA and define the host's susceptibility to certain diseases. High throughput sequencing of RNA will allow for more precise expression monitoring via transcriptomics.

Realistically speaking though, we currently dispose of two main techniques for distinction between MSSA and MRSA: those targeting specific diagnostic signature sequences and those that characterize entire chromosomes and then depend on bio-informatic analyses to highlight the presence of the same sequence motifs used by the specific methods (**Figure 1**).

TABLE 1 | Global review of future and commercial PCR tests for meticillin-resistant and -susceptible strains of *Staphylococcus aureus*.

Company	Status	Concise product description	Duration of test
Abacus Diagnostica, Finland	In development	Rapid DNA testing with proprietary GenomEra CDX-technology for identification of MRSA	50 min
AdvanDx, United States	FDA approved	*Staphylococcus* QuickFISH filter *in situ* hybridization test for positive blood culture liquid	20 min
Akonni Biosystems, United States	In development	TruArray MRSA, qualitative test for detection of SA and MRSA	Non-specified
Atlas Genetics, United Kingdom	In development	Mixed technology linking NAT and immunology for MRSA, Dual MRSA/MSSA	0.5 h
Autoi/mmun Diagnostika, Germany	CE certified	Automated AID Scanner, line probe Western blot probe assay after PCR amplification, 100 strips per hour	4 h
Biocartis, Belgium	In development	Idylla platform for multiplex real-time PCR assay for rapid detection of bloodstream infections	2 h
BioFire, United States	FDA approved, new tests in development	FDA approved syndromic panels for respiratory, gastro-intestinal, and meningitis/encephalitis associated pathogens; the BCID test also covers *mec*A. Sample in—result out strategy	1 h
BD, United States	FDA approved, new tests in development	Platform BD Max. MRSA + MSSA + *mec*A test	<3 h
Cepheid, United States (acquired by Danaher)	FDA approved for HAI with MRSA/SA	Validated for positive blood culture. Xpert test format. MRSA, SA Nasal Complete, MRSA/SA SSTI, MRSA/SA BC	2 h
Coyote Biosci, United States, China	In development	Platform Mini 8 RT PCR; throat swab/Blood sample—MRSA	10–30 min
Curetis AG, Germany	CE marked, precise status not very clear	Platform Univero; >100 pathogens and resistance genes, P55 Application focuses on pneumonia, 21 pathogens, and 19 resistance markers, 40-plex. i60 ITI Application Cartridge (23 organisms and 19 resistance genes)	4–5 h
DXna, United States	CE marked	GeneSTAT portable RT PCR platform, MRSA/MRCoNS in development for 2017	1 h
Epoch Biosciences, Elitech Group	FDA approved	Triplex Real Time Amplification tests using minor groove binding DNA probes	1 h
Genesig	RUO	Quantitative PCR for various targets among which MRSA; 16 samples per run	90–120 min
GenMark, United States	In development	Platform ePlex. Electronic sensor technology, DNA hybridization, and electrochemical detection	4 h
Genspeed, Austria	In development	Straightforward PCR with hybridization confirmation, combination of microfluidics, miniaturized opto-electronics, and automation	100 min
GFC Diagnostics	In development	Microscreen enzymatic-colorigenic DNA hybridization test on Safetube device	Non-specified
Great Basin Scientific, United States	Early stage	Whole blood, multiplexed nucleic-acid based assay using an opto-fluidic device; announced for 2021	Non-specified
Grenier Bio-One, United States	CE marked, not FDA cleared	PCR-based chip-probe Genspeed platform. Genspeed MRSA distinguishes MRSA/MRSE or *mec*A/C positive *S. haemolyticus*	1.5 h
Hain, Germany	CE marked for many tests	PCR/hybridization platform. GenoType, FluoroType and GenoQuick technologies, MRSA, CoNS	2.5 h
Icubate, United States	RUO	Random access multiplex PCR disposable test cassette for pathogens and resistances. Portfolio: gram + MSSA, *S. epidermidis*, MRSA	Non-specified
ID Biomedical, Corp., Vancouver	Early stage	Velogene rapid MRSA identification assay	2 h
Linear Diagnostics, Ltd.	In development	Detection of aligned substrate or PCR fragment via polarized light	Non-specified
Magnomics, Portugal	In development	Chip DNA extraction, amplification, and magnetic detection. Primary for veterinary application	1 h
Mobidiag, Finland	CE marked	Novodiag and Amplidiag product line. Sepsis, 60 bacterial species, 13 fungi, and *mec*A in one assay	3.5 h
Nanosphere Inc, United States	FDA cleared	DNA amplification-hybridization. Verigene BC-GP and BC-GN. Gold Nanoparticle Technology with oligo-hybridization to target DNA, narrow temperature range	2–2.5 h

(Continued)

Company	Status	Concise product description	Duration of test
Opgen, United States	Not clear	Real-time *C. difficile* and MRSA DNA testing. The Acuitas CRE Elite MDRO Gene Test detects 10 genes unique to MDRO from swabs to colonies	Non-specified
Pathogenica, Japan	RUO	Pathogenica's DxSeq is a sequencing product. The HAI BioDetection Kit assays over 12 pathogens and 15 resistance gene families	Non-specified
Roche Molecular Systems, Germany	FDA approved	COBAS 4800 and the Liat MRSA/MSSA PCR tests	1 h

Tests and companies may be missing and for several tests there are no peer-reviewed performance data. Diasorin and Luminex announced assays last year but the descriptions are not very precise. Same for the Spanish company Stat DX, which was recently acquired by Qiagen. Bruker and Siemens are said to be working on high throughput systems. Visibility of Chinese and Indian competitors in this field is also limited. Any mistakes are the authors' responsibility.
(BC)ID, (blood culture) identification; CE, Certification Europeènne; CoNS, coagulase negative staphylococci; FDA, Food and Drug Administration; GP/N, gram positive/negative; HAI, hospital acquired infection; MDRO, multi-drug resistant organism; (MR)SA, (methicillin-resistant) S. aureus; SSTI, skin and soft tissue infection.

TARGETED GENETIC DETECTION OF MSSA AND MRSA

It needs to be realized that for the molecular detection of MRSA there should be a differentiation between screening tests (for carriage) and hard-core diagnostic tests for infection (Osiecki, 2010; Trouillet-Assant et al., 2013). Both tests have different requirements for sensitivity, specificity, costs, and speed. Screening may not need high-speed but must be focused on specific detection of high-rate carriers (Bode et al., 2010). A major hurdle to developing molecular MRSA-specific tests is the fact that the gene encoding meticillin resistance occurs in other staphylococcal species as well. A solution to this issue is the inclusion of species-specific assays in the amplification reaction (e.g., targeting the *nuc* or *fem*A genes). The second hurdle is that the *mec* gene is present in four variants (*mec*A, *mec*C, and the more recently discovered *mec*B (Gómez-Sanz et al., 2015; Becker et al., 2018) and *mec*D (Schwendener et al., 2017)) and these reside in a growing number of genetic islands (Kolenda et al., 2017); *mec* genes are embedded in various Staphylococcal Cassette Chromosome *mec* (SCC*mec*) for which more than 10 different types have already been identified (Hill-Cawthorne et al., 2014; Kaya et al., 2018). Molecular tests for rapid discrimination of SCC*mec* types continue to be developed though their use is mostly of epidemiological rather than clinical value (Brukner et al., 2013). The *mec*C variant, mostly found in livestock associated MRSA and sharing about 70% sequence homology with *mec*A, was discovered as recently as 2011 (García-Álvarez et al., 2011). Finally, rapidity of a test can be affected by whether or not a test is (semi-)quantitative and performed in real-time or not (Verhoeven et al., 2012). The need for real-time testing differs per clinical application but in case of sepsis detection, for instance, speed is of utmost importance (Frye et al., 2012). Modern MRSA detection obviously is part of multiplexed, syndrome-oriented diagnostic testing (Blaschke et al., 2012; Ramanan et al., 2017).

A variety of experimental testing formats has been suggested for targeted MRSA detection but most of which have not reached the diagnostic market (yet). **Table 1** reviews the status of a significant number of current PCR tests and highlights a domain of importance: the next generation routine-applicable tests may very well-originate from this pool of potentially high throughput tools.

Some of the technologies are worth mentioning separately given the fact that they can be considered extremely elegant from an experimental design point of view. Digital droplet PCR for instance was shown to be sensitive and rapid and it has to be realized that instruments allowing in house development of droplet-based PCR tests are already (commercially) available (Luo et al., 2017). Nanowires are attractive because of size, relatively low costs, and speed of the assay and broad applicability of the technology which, in addition, is easy to multiplex (Ibarlucea et al., 2017). Similarly, using albumin stabilized fluorescent gold nanoclusters as selective probes, MSSA and MRSA can be reliably distinguished (Chan and Chen, 2012). Sensitivity and specificity of these often still quite experimental tests are usually good and offer a positive perspective on future developments in this field (Bakthavatchalam et al., 2017). Note that **Table 1** highlights the post-PCR use of array technology, filter *in situ* hybridization, minor groove binding DNA probes, and magnetic capturing as additional clever read-out methods.

LABORATORY-BASED vs. PoC TESTING

Classical testing for microbial pathogens usually leads to amplification of viable cells. This requires the use of specialized laboratories where employees and the community outside of the lab are protected from infection through specific control measures. This has often blocked PoC test development and deployment. Now, with the possibility to detect pathogens by amplifying non-infectious components of such pathogens the door toward out-of-laboratory testing has been opened wide. Miniaturized tools have been developed that are based on microfluidics (Yeh et al., 2017), LAMP combined with cellulose-based nucleic acid binding paper (Bearinger et al., 2011), isothermal amplification tests (Toley et al., 2015) but also based upon labeling- and amplification-free techniques (Corrigan et al., 2013). With such technologies in mind it was established that PoC testing for MSSA and MRSA was among the priorities when remote and even

disaster testing was due (Brock et al., 2010; Kost et al., 2012). Clearly, tools for bedside diagnostics are available that allow for in-department infection control and outbreak management.

In PoC testing both technical and clinical aspects are of key importance. Technical requirements are largely covered by the WHO ASSURED criteria. The acronym lists affordability, sensitivity, specificity, user-friendliness, rapidity and robustness, no need for complicated equipment, and providing solutions that can be easily delivered to end users. If all these requirement are met in a single test (which at this stage is non-existent) then clinical applicability is essentially global. However, if a test would, for example, be too expensive then application in developing economies would essentially be blocked. MRSA/MSSA PoC tests would be particularly useful for rapid assessment of (nasal) carriage for infection control, whereas screening for staphylococcal wound infection and respiratory infection would also have strong added value.

The first PoC MRSA projects have been published. Leone et al. (2013) did an intensive care-based study into the use of MSSA/MRSA detection in patients with ventilator-associated pneumonia. They showed that with a negative predictive value of 99.8% PoC testing efficiently excluded the presence of MRSA among the patients. They did warn that the reliability of this type of testing is dependent on the local prevalence of MRSA carriage. In an orthopedic readmission study it was shown that the Cepheid Xpert MRSA with its 75% sensitivity in this groups of patients with complicated problems performed quite well (Parcell and Phillips, 2014). Screening more than 10,000 patients at admission for detection of MRSA carriage was very efficient as well (Wu et al., 2017). Although the list of publications is relatively short, it is clear that detection of MSSA/MRSA at the PoC fulfills a real medical need. In case of epidemic spread of MRSA clones rapid and sensitive detection are key and in many cases the use of PoC testing allows for accelerated testing in comparison with more conventional laboratory assays. Speed really is the key to high throughput surveillance and subsequent rapid infection control.

GENOMIC DETECTION OF MSSA AND MRSA

Single genomic molecules can be captured to microscopic beads, which are equipped with biotinylated probes to which streptavidin-complexed galactosidase binds and which facilitates the detection of sub-femtomolar concentrations of specific DNA molecules. This Single Molecule Array tests (developed by Quanterix Corporation, Lexington, MA, United States) has been adapted for the detection of MRSA as well (Song et al., 2013). Beyond capturing and detecting "full" genomes, there is now also the option to have a staphylococcal genome sequenced *de novo* and *in toto*. Thousands of MSSA and MRSA strains have been subjected to genome sequencing and the software that allows for post-sequencing detection of the *mec* genes is available (Gordon et al., 2014). Current and well-known software packages for such purposes include CLC Bio (Qiagen,

Hilden, Germany), Seqsphere (Ridom, Münster, Germany), and Bionumerics (Applied Maths, bioMérieux, St. Martens Lathem, Belgium). Hence genomic characterization of MRSA is feasible and with the rise in sequencing directly from clinical specimens the impact of direct detection of MRSA will change significantly over the years to come (Lefterova et al., 2015). However, in order to be applicable in routine high-throughput clinical laboratories the technology needs to be quicker, less expensive with data that should be easy to interpret preferably in a (semi-)quantitative fashion. Obviously, genome sequencing provides the ultimate tool for epidemiological typing of MRSA and MSSA (Quainoo et al., 2017) and many studies where NGS has been exploited to define epidemiological patterns of spreading of MRSA have been published before (Tewhey et al., 2012; Price et al., 2013; Stone et al., 2013; Kong et al., 2016; Ward et al., 2016; Planet et al., 2017; and references therein).

CONCLUDING REMARKS

Whereas classical detection and speciation of staphylococci has improved significantly upon the introduction of MALDI ToF MS in the diagnostic laboratory, molecular tests, mostly based on specific gene amplification, are still required for the rapid distinction between MSSA and MRSA. The availability of WGS and NGS has now opened up alternative avenues for the detection of resistance genes, the *mec*-variants included. The near future will bring (genome) sequencing and comprehensive software packages allowing for the unequivocal bio-informatic AST of MSSA and MRSA using WGS, even for non-bio-informaticians. The position of PoC testing in all of this is still poorly defined and needs to be clarified. Inclusion of additional patient data beyond laboratory results is an important additive to the optimization of PoC testing (Yoshioka et al., 2018). In conclusion, molecular testing for MRSA has been accepted by the diagnostic community and is performing well. New technology will challenge the molecular tests and there will be fierce clinical, commercial, and academic completion before full acceptation of the new wave of genomic testing formats.

AUTHOR CONTRIBUTIONS

AB and OR: conceived, wrote and illustrated the manuscript. Data for the table was assembled using aimed searches in relevant databases and publicly available information in a variety of corporate websites and communications between company employees and the authors. Omissions in the table are for the responsibility of the authors only.

FUNDING

REFERENCES

Ábrók, M., Lázár, A., Szécsényi, M., Deák, J., and Urbán, E. (2018). Combination of MALDI-TOF MS and PBP2' latex agglutination assay for rapid MRSA detection. *J. Microbiol. Methods* 144, 122–124. doi: 10.1016/j.mimet.2017.11.021

Almassian, D. R., Cockrell, L. M., and Nelson, W. M. (2013). Portable nucleic acid thermocyclers. *Chem. Soc. Rev.* 42, 8769–8798. doi: 10.1039/c3cs60144g

Bakthavatchalam, Y. D., Nabarro, L. E., and Veeraraghavan, B. (2017). Evolving rapid methicillin-resistant *Staphylococcus aureus* detection: cover all the bases. *J. Glob. Infect. Dis.* 9, 18–22. doi: 10.4103/0974-777X.199997

Bearinger, J. P., Dugan, L. C., Baker, B. R., Hall, S. B., Ebert, K., Mioulet, V., et al. (2011). Development and initial results of a low cost, disposable, point-of-care testing device for pathogen detection. *IEEE Trans. Biomed. Eng.* 58, 805–808. doi: 10.1109/TBME.2010.2089054

Becker, K., van Alen, S., Idelevich, E. A., Schleimer, N., Seggewiß, J., Mellmann, A., et al. (2018). Plasmid-encoded transferable mecB-mediated methicillin resistance in *Staphylococcus aureus*. *Emerg. Infect. Dis.* 24, 242–248. doi: 10.3201/eid2402.171074

Bernardo, K., Pakulat, N., Macht, M., Krut, O., Seifert, H., Fleer, S., et al. (2002). Identification and discrimination of *Staphylococcus aureus* strains using matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Proteomics* 2, 747–753. doi: 10.1002/1615-9861(200206)2:6<747::AID-PROT747>3.0.CO;2-V

Blaschke, A. J., Heyrend, C., Byington, C. L., Fisher, M. A., Barker, E., Garrone, N. F., et al. (2012). Rapid identification of pathogens from positive blood cultures by multiplex polymerase chain reaction using the FilmArray system. *Diagn. Microbiol. Infect. Dis.* 74, 349–355. doi: 10.1016/j.diagmicrobio.2012.08.013

Bode, L. G., Kluytmans, J. A., Wertheim, H. F., Bogaers, D., Vandenbroucke-Grauls, C. M., Roosendaal, R., et al. (2010). Preventing surgical-site infections in nasal carriers of *Staphylococcus aureus*. *N. Engl. J. Med.* 362, 9–17. doi: 10.1056/NEJMoa0808939

Brock, T. K., Mecozzi, D. M., Sumner, S., and Kost, G. J. (2010). Evidence-based point-of-care tests and device designs for disaster preparedness. *Am. J. Disaster Med.* 5, 285–294.

Brown, J., and Paladino, J. A. (2010). Impact of rapid methicillin-resistant *Staphylococcus aureus* polymerase chain reaction testing on mortality and cost effectiveness in hospitalized patients with bacteraemia: a decision model. *Pharmacoeconomics* 28, 567–575. doi: 10.2165/11533020-000000000-00000

Brukner, I., Oughton, M., Giannakakis, A., Kerzner, R., and Dascal, A. (2013). Significantly improved performance of a multitarget assay over a commercial SCCmec-based assay for methicillin-resistant *Staphylococcus aureus* screening: applicability for clinical laboratories. *J. Mol. Diagn.* 15, 577–580. doi: 10.1016/j.jmoldx.2013.04.009

Camoez, M., Sierra, J. M., Dominguez, M. A., Ferrer-Navarro, M., Vila, J., and Roca, I. (2016). Automated categorization of methicillin-resistant *Staphylococcus aureus* clinical isolates into different clonal complexes by MALDI-TOF mass spectrometry. *Clin. Microbiol. Infect.* 22, .e1–.e161. doi: 10.1016/j.cmi.2015.10.009

Chan, P. H., and Chen, Y. C. (2012). Human serum albumin stabilized gold nanoclusters as selective luminescent probes for *Staphylococcus aureus* and methicillin-resistant *Staphylococcus aureus*. *Anal. Chem.* 84, 8952–8956. doi: 10.1021/ac302417k

Charretier, Y., Dauwalder, O., Franceschi, C., Degout-Charmette, E., Zambardi, G., Cecchini, T., et al. (2015). Rapid bacterial identification, resistance, virulence and type profiling using selected reaction monitoring mass spectrometry. *Sci. Rep.* 5:13944. doi: 10.1038/srep13944

Corrigan, D. K., Schulze, H., Henihan, G., Hardie, A., Ciani, I., Giraud, G., et al. (2013). Development of a PCR-free electrochemical point of care test for clinical detection of methicillin resistant *Staphylococcus aureus* (MRSA). *Analyst* 138, 6997–7005. doi: 10.1039/c3an01319g

Du, Z., Yang, R., Guo, Z., Song, Y., and Wang, J. (2002). Identification of *Staphylococcus aureus* and determination of its methicillin resistance by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* 74, 5487–5491. doi: 10.1021/ac020109k

Frye, A. M., Baker, C. A., Rustvold, D. L., Heath, K. A., Hunt, J., Leggett, J. E., et al. (2012). Clinical impact of a real-time PCR assay for rapid identification of staphylococcal bacteremia. *J. Clin. Microbiol.* 50, 127–133. doi: 10.1128/JCM.06169-11

García-Álvarez, L., Holden, M. T., Lindsay, H., Webb, C. R., Brown, D. F., Curran, M. D., et al. (2011). Meticillin-resistant *Staphylococcus aureus* with a novel mecA homologue in human and bovine populations in the UK and Denmark: a descriptive study. *Lancet Infect. Dis.* 11, 595–603. doi: 10.1016/S1473-3099(11)70126-8

Goldstein, J. E., Zhang, L., Borror, C. M., Rago, J. V., and Sandrin, T. R. (2013). Culture conditions and sample preparation methods affect spectrum quality and reproducibility during profiling of *Staphylococcus aureus* with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Lett. Appl. Microbiol.* 57, 144–150. doi: 10.1111/lam.12092

Gómez-Sanz, E., Schwendener, S., Thomann, A., Gobeli Brawand, S., and Perreten, V. (2015). First Staphylococcal cassette chromosome *mec* containing a *mecB*-carrying gene complex independent of transposon Tn*6045* in a *Macrococcus caseolyticus* isolate from a canine infection. *Antimicrob. Agents Chemother.* 59, 4577–4583. doi: 10.1128/AAC.05064-14

Gootenberg, J. S., Abudayyeh, O. O., Kellner, M. J., Joung, J., Collins, J. J., and Zhang, F. (2018). Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* 360, 439–444. doi: 10.1126/science.aaq0179

Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Essletzbichler, P., Dy, A. J., Joung, J., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356, 438–442. doi: 10.1126/science.aam9321

Gordon, N. C., Price, J. R., Cole, K., Everitt, R., Morgan, M., Finney, J., et al. (2014). Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing. *J. Clin. Microbiol.* 52, 1182–1191. doi: 10.1128/JCM.03117-13

Henares, D., Brotons, P., Buyse, X., Latorre, I., de Paz, H. D., and Muñoz-Almagro, C. (2017). Evaluation of the eazyplex MRSA assay for the rapid detection of *Staphylococcus aureus* in pleural and synovial fluid. *Int. J. Infect. Dis.* 59, 65–68. doi: 10.1016/j.ijid.2017.04.013

Hill-Cawthorne, G. A., Hudson, L. O., El Ghany, M. F., Piepenburg, O., Nair, M., Dodgson, A., et al. (2014). Recombinations in staphylococcal cassette chromosome *mec* elements compromise the molecular detection of methicillin resistance in *Staphylococcus aureus*. *PLoS One* 9:e101419. doi: 10.1371/journal.pone.0101419

Ibarlucea, B., Rim, T., Baek, C. K., de Visser, J. A. G. M., Baraban, L., and Cuniberti, G. (2017). Nanowire sensors monitor bacterial growth kinetics and response to antibiotics. *Lab Chip* 17, 4283–4293. doi: 10.1039/c7lc00807d

Kaya, H., Hasman, H., Larsen, J., Stegger, M., Johannesen, T. B., Allesøe, R. L., et al. (2018). SCC*mec*Finder, a web-based tool for typing of Staphylococcal cassette chromosome mec in *Staphylococcus aureus* using whole-genome sequence data. *mSphere* 3:e00612-17. doi: 10.1128/mSphere.00612-17

Kelly, R., Zoubiane, G., Walsh, D., Ward, R., and Goossens, H. (2016). Public funding for research on antibacterial resistance in the JPIAMR countries, the European Commission, and related European Union agencies: a systematic observational analysis. *Lancet Infect. Dis.* 16, 431–440. doi: 10.1016/S1473-3099(15)00350-3

Kolenda, C., Dupieux, C., Decousser, J. W., Larsen, A. R., Pichon, B., Holmes, M., et al. (2017). Comparison of automated antimicrobial susceptibility testing systems to detect *mecC*-positive methicillin-resistant *Staphylococcus aureus*. *J. Clin. Microbiol.* 55, 3554–3556. doi: 10.1128/JCM.01150-17

Kong, H., Tong, L., Zhang, W., Fu, Y., and Li, X. (2014). Combined use of the BinaxNOW *Staphylococcus aureus* test with the Clearview PBP2a assay for the early detection of methicillin-resistant S. *aureus* from positive blood cultures. *Diagn. Microbiol. Infect. Dis.* 78, 226–228. doi: 10.1016/j.diagmicrobio.2013.11.018

Kong, Z., Zhao, P., Liu, H., Yu, X., Qin, Y., Su, Z., et al. (2016). Whole-genome sequencing for the investigation of a hospital outbreak of MRSA in China. *PLoS One* 11:e0149844. doi: 10.1371/journal.pone.0149844

Kost, G. J., Mecozzi, D. M., Brock, T. K., and Curtis, C. M. (2012). Assessing point-of-care device specifications and needs for pathogen detection in emergencies and disasters. *Point Care* 11, 119–125. doi: 10.1097/POC.0b013e31825a25cb

Lefterova, M. I., Suarez, C. J., Banaei, N., and Pinsky, B. A. (2015). Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. *J. Mol. Diagn.* 17, 623–634. doi: 10.1016/j.jmoldx.2015.07.004

Leone, M., Malavieille, F., Papazian, L., Meyssignac, B., Cassir, N., Textoris, J., et al. (2013). Routine use of *Staphylococcus aureus* rapid diagnostic test in

patients with suspected ventilator-associated pneumonia. *Crit. Care* 17:R170. doi: 10.1186/cc12849

Liu, C., Bayer, A., Cosgrove, S. E., Daum, R. S., Fridkin, S. K., Gorwitz, R. J., et al. (2011). Clinical practice guidelines by the infectious diseases society of America for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children. *Clin. Infect. Dis.* 52, e18–e55. doi: 10.1093/cid/ciq146

Luo, J., Li, J., Yang, H., Yu, J., and Wei, H. (2017). Accurate detection of methicillin-resistant *Staphylococcus aureus* in mixtures by use of single-bacterium duplex droplet digital PCR. *J. Clin. Microbiol.* 55, 2946–2955. doi: 10.1128/JCM.00716-17

Lüthje, P., Pranada, A. B., Carruthers-Lay, D., Desjardins, M., Gaillot, O., Wareham, D., et al. (2017). Identification of microorganisms grown on chromogenic media by MALDI-TOF MS. *J. Microbiol. Methods* 136, 17–20. doi: 10.1016/j.mimet.2017.03.001

Miyamoto, M., Motooka, D., Gotoh, K., Imai, T., Yoshitake, K., Goto, N., et al. (2014). Performance comparison of second- and third-generation sequencers using a bacterial genome with two chromosomes. *BMC Genomics* 15:699. doi: 10.1186/1471-2164-15-699

Muldrew, K. L. (2009). Molecular diagnostics of infectious diseases. *Curr. Opin. Pediatr.* 21, 102–111. doi: 10.1097/MOP.0b013e328320d87e

Okeke, I. N., Peeling, R. W., Goossens, H., Auckenthaler, R., Olmsted, S. S., de Lavison, J. F., et al. (2011). Diagnostics as essential tools for containing antibacterial resistance. *Drug Resist. Updat.* 14, 95–106. doi: 10.1016/j.drup.2011.02.002

Osiecki, J. (2010). *MRSA Screening: Culture or Molecular?* Available at: www.advanceweb.com

Østergaard, C., Hansen, S. G., and Møller, J. K. (2015). Rapid first-line discrimination of methicillin resistant *Staphylococcus aureus* strains using MALDI-TOF MS. *Int. J. Med. Microbiol.* 305, 838–847. doi: 10.1016/j.ijmm.2015.08.002

Parcell, B. J., and Phillips, G. (2014). Use of Xpert MRSA PCR point-of-care testing beyond the laboratory. *J. Hosp. Infect.* 87, 119–121. doi: 10.1016/j.jhin.2014.04.002

Perry, J. D. (2017). A decade of development of chromogenic culture media for clinical microbiology in an era of molecular diagnostics. *Clin. Microbiol. Rev.* 30, 449–479. doi: 10.1128/CMR.00097-16

Planet, P. J., Narechania, A., Chen, L., Mathema, B., Boundy, S., Archer, G., et al. (2017). Architecture of a species: phylogenomics of *Staphylococcus aureus*. *Trends Microbiol.* 25, 153–166. doi: 10.1016/j.tim.2016.09.009

Price, J., Gordon, N. C., Crook, D., Llewelyn, M., and Paul, J. (2013). The usefulness of whole genome sequencing in the management of *Staphylococcus aureus* infections. *Clin. Microbiol. Infect.* 19, 784–789. doi: 10.1111/1469-0691.12109

Quainoo, S., Coolen, J. P. M., van Hijum, S. A. F. T., Huynen, M. A., Melchers, W. J. G., van Schaik, W., et al. (2017). Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin. Microbiol. Rev.* 30, 1015–1063. doi: 10.1128/CMR.00016-17

Ramanan, P., Bryson, A. L., Binnicker, M. J., Pritt, B. S., and Patel, R. (2017). Syndromic panel-based testing in clinical microbiology. *Clin. Microbiol. Rev.* 31:e00024-17. doi: 10.1128/CMR.00024-17

Rees, J. C., and Barr, J. R. (2017). Detection of methicillin-resistant *Staphylococcus aureus* using phage amplification combined with matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Bioanal. Chem.* 409, 1379–1386. doi: 10.1007/s00216-016-0070-3

Rhoads, D. D., Wang, H., Karichu, J., and Richter, S. S. (2016). The presence of a single MALDI-TOF mass spectral peak predicts methicillin resistance in staphylococci. *Diagn. Microbiol. Infect. Dis.* 86, 257–261. doi: 10.1016/j.diagmicrobio.2016.08.001

Rossen, J. W. A., Friedrich, A. W., Moran-Gilad, J. and ESCMID study group for genomic and molecular diagnostics (ESGMD) (2017). Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin. Microbiol. Infect.* 24, 355–360. doi: 10.1016/j.cmi.2017.11.001

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Sci. Rep.* 7:43797. doi: 10.1038/srep43797

Sogawa, K., Watanabe, M., Ishige, T., Segawa, S., Miyabe, A., Murata, S., et al. (2017). Rapid discrimination between methicillin-sensitive and methicillin-resistant *Staphylococcus aureus* using MALDI-TOF mass spectrometry. *Biocontrol Sci.* 22, 163–169. doi: 10.4265/bio.22.163

Song, L., Shan, D., Zhao, M., Pink, B. A., Minnehan, K. A., York, L., et al. (2013). Direct detection of bacterial genomic DNA at sub-femtomolar concentrations using single molecule arrays. *Anal. Chem.* 85, 1932–1939. doi: 10.1021/ac303426b

Stone, M. J., Wain, J., Ivens, A., Feltwell, T., Kearns, A. M., and Bamford, K. B. (2013). Harnessing the genome: development of a hierarchical typing scheme for meticillin-resistant *Staphylococcus aureus*. *J. Med. Microbiol.* 62(Pt 1), 36–45. doi: 10.1099/jmm.0.049957-0

Suh, H. J., Park, W. B., Jung, S.-I., Song, K.-H., Kwak, Y. G., Kim, K.-H., et al. (2018). A risk scoring system for predicting methicillin resistance in community-onset *Staphylococcus aureus* bacteremia in Korea. *Microb. Drug Resist.* 24, 556–562. doi: 10.1089/mdr.2017.0236

Szabados, F., Woloszyn, J., Richter, C., Kaase, M., and Gatermann, S. (2010). Identification of molecularly defined *Staphylococcus aureus* strains using matrix-assisted laser desorption/ionization time of flight mass spectrometry and the Biotyper 2.0 database. *J. Med. Microbiol.* 59(Pt 7), 787–790. doi: 10.1099/jmm.0.016733-0

Tewhey, R., Cannavino, C. R., Leake, J. A., Bansal, V., Topol, E. J., Torkamani, A., et al. (2012). Genetic structure of community acquired methicillin-resistant *Staphylococcus aureus* USA300. *BMC Genomics* 13:508. doi: 10.1186/1471-2164-13-508

Toley, B. J., Covelli, I., Belousov, Y., Ramachandran, S., Kline, E., Scarr, N., et al. (2015). Isothermal strand displacement amplification (iSDA): a rapid and sensitive method of nucleic acid amplification for point-of-care diagnosis. *Analyst* 140, 7540–7549. doi: 10.1039/c5an01632k

Trouillet-Assant, S., Rasigade, J. P., Lustig, S., Lhoste, Y., Valour, F., Guerin, C., et al. (2013). Ward-specific rates of nasal cocolonization with methicillin-susceptible and -resistant *Staphylococcus* spp. and potential impact on molecular methicillin-resistant *Staphylococcus aureus* screening tests. *J. Clin. Microbiol.* 51, 2418–2420. doi: 10.1128/JCM.00491-13

Turnidge, J. D., Kotsanas, D., Munckhof, W., Roberts, S., Bennett, C. M., Nimmo, G. R., et al. (2009). *Staphylococcus aureus* bacteraemia: a major cause of mortality in Australia and New Zealand. *Med. J. Aust.* 191, 368–373.

Ueda, O., Tanaka, S., Nagasawa, Z., Hanaki, H., Shobuike, T., and Miyamoto, H. (2015). Development of a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrum (MALDI-TOF-MS)-based typing method to identify meticillin-resistant *Staphylococcus aureus* clones. *J. Hosp. Infect.* 90, 147–155. doi: 10.1016/j.jhin.2014.11.025

Verhoeven, P. O., Grattard, F., Carricajo, A., Lucht, F., Cazorla, C., Garraud, O., et al. (2012). Quantification by real-time PCR assay of *Staphylococcus aureus* load: a useful tool for rapidly identifying persistent nasal carriers. *J. Clin. Microbiol.* 50, 2063–2065. doi: 10.1128/JCM.00157-12

Verroken, A., Defourny, L., le Polain de Waroux, O., Belkhir, L., Laterre, P. F., Delmée, M., et al. (2016). Clinical impact of MALDI-TOF MS identification and rapid susceptibility testing on adequate antimicrobial treatment in sepsis with positive blood cultures. *PLoS One* 11:e0156299. doi: 10.1371/journal.pone.0156299

Ward, M. J., Goncheva, M., Richardson, E., McAdam, P. R., Raftis, E., Kearns, A., et al. (2016). Identification of source and sink populations for the emergence and global spread of the East-Asia clone of community-associated MRSA. *Genome Biol.* 17:160. doi: 10.1186/s13059-016-1022-0

Wu, P. J., Jeyaratnam, D., Tosas, O., Cooper, B. S., and French, G. L. (2017). Point-of-care universal screening for meticillin-resistant *Staphylococcus aureus*: a cluster-randomized cross-over trial. *J. Hosp. Infect.* 95, 245–252. doi: 10.1016/j.jhin.2016.08.017

Yan, L., Zhou, J., Zheng, Y., Gamson, A. S., Roembke, B. T., Nakayama, S., et al. (2014). Isothermal amplified detection of DNA and RNA. *Mol. Biosyst.* 10, 970–1003. doi: 10.1039/c3mb70304e

Yeh, E. C., Fu, C. C., Hu, L., Thakur, R., Feng, J., and Lee, L. P. (2017). Self-powered integrated microfluidic point-of-care low-cost enabling (SIMPLE) chip. *Sci. Adv.* 3:e1501645. doi: 10.1126/sciadv.1501645

Yoshioka, N., Deguchi, M., Hagiya, H., Yoshida, H., Yamamoto, N., Hashimoto, S., et al. (2018). Available, bed-sided, comprehensive (ABC) score to a diagnosis

of Methicillin-resistant *Staphylococcus aureus* infection: a derivation and validation study. *BMC Infect. Dis.* 18:19. doi: 10.1186/s12879-017-2919-2

Zhang, T., Ding, J., Rao, X., Yu, J., Chu, M., Ren, W., et al. (2015). Analysis of methicillin-resistant *Staphylococcus aureus* major clonal lineages by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS). *J. Microbiol. Methods* 117, 122–127. doi: 10.1016/j.mimet.2015.08.002

Zhu, W., Sieradzki, K., Albrecht, V., McAllister, S., Lin, W., Stuchlik, O., et al. (2015). Evaluation of the Biotyper MALDI-TOF MS system for identification of *Staphylococcus* species. *J. Microbiol. Methods* 117, 14–17. doi: 10.1016/j.mimet.2015.07.014

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership