S. epidermidis CELL WALL ANCHORED PROTEINS LOCATED IN COMPOSITE SCC*mec* ISLANDS

A Dissertation

by

SRISHTEE ARORA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Magnus Hook
Committee Members,	Yi Xu
	Jon Skare
	Peter Davies
	Samuel Shelburne
Head of Program,	Warren Zimmer

December 2017

Major Subject: Medical Sciences

Copyright 2017 Srishtee Arora

ABSTRACT

Pathogenic bacteria improve their fitness, pathogenicity and virulence by acquisition of mobile genetic elements (MGE). These MGE carry genes for antibiotic resistance, toxins, virulence factors, etc. *Staphylococcus epidermidis* is a leading cause of nosocomial infections in patients with a compromised immune system and an implanted medical device. It has an open pan-genome with 20% variable genes. In my studies of *S. epidermidis* clinical isolates, I have discovered two genes that code for surface proteins, *S. epidermidis* surface protein J (SesJ) and Plasmin sensitive protein (Pls). These genes are present in a Staphylococcal Cassette Chromosome containing methicillin resistance gene (SCC*mec*). Further analysis has shown that SesJ contains structural features characteristic of Microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), known major virulence factors of Gram-positive pathogenic bacteria. SesJ also contains a N-terminal repeat (NTR) region that has not been previously observed in MSCRAMMs. SesJ structural homologs were identified in three other coagulase negative staphylococci composing a new MSCRAMM NTR-containing subfamily.

171 clinical isolates were retrospectively collected from MD Anderson, Houston, and examined for the presence and location of the *sesJ* and the *pls* genes. Work presented in this dissertation show that the *sesJ* gene is prevalent in ST2, ST5 and ST 210 isolates, while the *pls* gene is present only in ST2. ST2 and ST5 isolates accounted for over 50%

of the total isolates. Genetic location identification showed that *sesJ* is present in SCC*mec* Type IV and SCC*mec* Type VII. Different isoforms of SesJ were also observed in CC2 and CC5 isolates. Pls has been previously reported in *S. aureus*, where it is present in the SCC*mec* Type I. However in this study, the *pls* gene was discovered in *S. epidermidis* SCC*mec* Type IV. Additionally, the frequency of the *sesJ* gene has increased three folds in past three years. In summary, I discovered two novel surface proteins present in SCC*mec*. In the future, we will corroborate the presence of these genes with clinical outcome in patients and determine the role of SesJ in the *S. epidermidis* pathogenicity.

DEDICATION

I dedicate this dissertation to my parents and brother. Thank you for all you have done for me! Thank you for supporting my crazy decisions like moving to a country far away where I knew no one. Even with so much distance between us, you all were always there for me through ups and downs of life.

ACKNOWLEDGEMENTS

First, I would like to thank my committee chair, Dr. Magnus Höök for his mentorship over the past 8.5 years. Thank you for giving me an opportunity to do an internship in your lab, which strengthened my belief to pursue a Ph.D. degree. You have taught me to be an independent researcher and a critical thinker. You always had an open door policy to discuss anything from science to personal life, which made the tough process of pursuing a Ph.D. easier. I deeply appreciate all your support during the rough times. I enjoy your questions “how is it going?” and “what am I discovering today?” even if I beat you to it some days. Importantly, I look forward to our future endeavors together!

I would also like to thank other members of my committee Dr. Jon Skare, Dr. Samuel Shelburne, Dr. Yi Xu, and Dr. Peter Davies, for their guidance and support throughout the course of this research. Dr. Shelburne, I appreciate that you made our collaboration open and easy. I am grateful for your insights into my project.

Mata, Pita ji and Bhai, I am so blessed to have you all as my family. Thank you for believing in me and supporting me through my journey. Mata, I know my decision to move to the US was especially hard on you. I cannot appreciate enough the support you have shown me through these years. Bhai, I wish we lived closer. Pita ji, thank you for understanding my goals in life were different than people around us.

Thanks also go to my friends Jose, Jerry, Danielle, Tora and Andrea. I would like to thank Jose for editing my dissertation. Jose and Jerry you provided me a family away from home. Danielle, I can never forget practicing my presentation with you in a restaurant. Thank you for always being available for my presentation rehearsals. Andrea I always enjoy our motivational talks. Thank you all my present and past colleagues at the center for making my time at Texas A&M University a great experience.

Finally, I would like to thank Mark for all his love, support and encouragement! Thank you for your gracious sacrifices of our weekend plans because of my weekend work schedule.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Dr. Magnus Hook [advisor] and Dr. Yi Xu, Dr. Peter Davies, Dr. Jon Skare and Dr. Samuel Shelburne.

Epidemiology data in Chapter II was provided by Dr. Anne-Catrin Uhlemann. The isolates used in Chapter III were provided by Dr. Samuel Shelburne. Library preparation for the clinical isolate was done by Dr. Samuel Shelburne's laboratory. Genome assembly of the isolates was done by Xiqi Li under supervision of Dr. Awdhesh Kalia and Dr. Samuel Shelburne.

All other work conducted for the dissertation was completed by the student - Srishtee Arora.

Funding Sources

This work was made possible in part by National Institutes of Health under Grant Number RO1 AI020624.

NOMENCLATURE

Aae	Autolysin/adhesin of *S.epidermidis*
Aap	Accumulation associated protein
ACME	Arginine catabolic mobile element
AltE	Autolysin E
Bbp	Bonesialo binding protein
CHIP	Chemotaxis inhibitory protein
CLABSI	Central line associated bloodstream infection
ClfA	Clumping factor A
ClfB	Clumping factor B
Cna	Collagen adhesion
CoNS	Coagulase negative *Staphylococcus*
CWA	Cell wall anchored
DLL	Dock lock and latch
DNA	Deoxyribonucleic acid
Ebp	Elastin binding protein
Embp	Extracellular matrix binding protein
Esp	*S. epidermidis* serine protease
FBR	Foreign body related
FBR-BSI	Foreign body related-bloodstream infection
FBRI	Foreign body related infection

FnBP	Fibronectin binding protein
GupSE	Glutamyl endopeptidase
HGT	Horizontal gene transfer
ICE	Integrated conjugative element
IgG	Immunoglobulin G
MGE	Mobile genetic element
MLS B	Macrolide, lincosamide and streptogramin B
MSCRAMM	Microbial surface components recognizing adhesive matrix Molecules
PBP	Penicillin binding protein
PBP2′	Penicillin binding proteins 2′
PGA	Poly-γ-DL-glutamic acid
PIA	Polysaccharide intercellular adhesion
Pls	Plasmin sensitive protein
PSM	Phenol soluble modulin
RNA	Ribonucleic acid
SasX	*Staphylococcus aureus* surface protein X
SCC	Staphylococcal cassette chromosome
SCIN	Staphylococcal inhibitor of complement
SdrC	Serine aspartate dipeptide repeat protein C
SdrD	Serine aspartate dipeptide repeat protein D
SdrE	Serine aspartate dipeptide repeat protein E

SdrF	Serine aspartate dipeptide repeat protein F
SdrG	Serine aspartate dipeptide repeat protein G
SEC3	Staphylococcal enterotoxin C3
SEIL	Staphylococcal enterotoxin-like toxin L
SepA	Staphylococcal efflux pump A
SePI	*Staphylococcus epidermidis* pathogenicity island
SesC	*Staphylococcus epidermidis* surface protein C
SesI	*Staphylococcus epidermidis* surface protein I
SesJ	*Staphylococcus epidermidis* surface protein J
SspA	Staphylococcal serine protease A
ST	Sequence type
tRNA	Transfer ribonucleic acid
TSST	Toxic shock syndrome toxin

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

CONTRIBUTORS AND FUNDING SOURCES vii

NOMENCLATURE viii

TABLE OF CONTENTS xi

LIST OF FIGURES xiii

LIST OF TABLES xiv

CHAPTER I INTRODUCTION AND LITERATURE REVIEW 1

Clinical relevance of the dissertation 1
S. epidermidis and other CoNS 2
MSCRAMMs 11
Mobile genetic elements 21
Horizontal gene transfer 37
CoNS as a reservoir of antibiotic resistance 39

CHAPTER II A NOVEL MSCRAMM SUBFAMILY IN COAGULASE NEGATIVE STAPHYLOCOCCAL SPECIES 41

Introduction 42
Materials and methods 44
Results 47
Discussion 62

CHAPTER III SESJ AND PLS ARE PRESENT IN A COMPOSITE SCC*mec* ELEMENT 68

Introduction 68
Materials and methods 71
Results 80
Discussion 90

CHAPTER IV CONCLUSIONS 95

REFERENCES 107

LIST OF FIGURES

Page

Figure 1. Schematic representation of biofilm formation ... 6

Figure 2. Cartoon representation of previously characterized subfamilies of MSCRAMM ... 13

Figure 3. Dock, lock and latch model ... 16

Figure 4. Collagen hug model ... 17

Figure 5. Structural organization of SCC*mec* ... 27

Figure 6. Schematic representation of SCC*mec* types ... 30

Figure 7. Synthesis of peptidoglycan for cell wall formation ... 32

Figure 8. The *mec* operon and its expression in absence/presence of methicillin ... 34

Figure 9. Integration and excision of SCC*mec* ... 36

Figure 10. Mechanisms of horizontal gene transfer in prokaryotes ... 38

Figure 11. Comparison of SesJ with other structurally related proteins ... 49

Figure 12. Detection of SesJ expression on the surface of *S. epidermidis* by flow cytometry ... 56

Figure 13. SesJ structural homologs in other CoNS species ... 59

Figure 14. Conserved motif in the NTRs of SesJ structural homologs ... 63

Figure 15. The *sesJ* gene is located in composite SCC*mec* element ... 86

Figure 16. Sequence comparison of SesJ A-region from different STs ... 87

Figure 17. Frequency of the *sesJ*$^{+}$ *S. epidermidis* isolates is increasing ... 91

LIST OF TABLES

Page

Table 1. Virulence factors of *S. epidermidis* 10

Table 2. Functions of unknown MSCRAMMs 19

Table 3. The *ccr* and *mec* gene complex present in SCC*mec* types 29

Table 4. Conserved motifs in the A-region of SesJ structural homologs 51

Table 5. Summary of NTRs of SesJ structural homologs 53

Table 6. Presence of virulence genes in the isolates 74

Table 7**.** Presence of the *sesJ* and the *pls* gene in different ST 81

Table 8. Presence of the *sesJ* and *pls* gene in MRSE and MSSE isolates 83

Table 9. SesJ isoforms are present in different clonal complexes 89

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Clinical relevance of the dissertation

Staphylococcal species are the leading cause of nosocomial infections in the USA. Methicillin resistant *Staphylococcus aureus* alone causes about 80,000 severe invasive infections per year and 11,285 of these infections lead to death in the U.S. alone [1]. Coagulase negative staphylococci (CoNS), especially *S. epidermidis* often causes infections in immunosuppressed/immunocompromised patients and patients with inserted/implanted foreign material. *S. epidermidis* is the leading causative agent of inserted/implanted foreign body-related infections (FBRIs), which usually result in foreign body – related bloodstream infections (FBR-BSIs). Patients with FBR-BSIs characteristically have longer hospital/ICU stays as well as higher mortality rates and associated costs compared to uninfected patients.

Staphylococci are notorious for developing resistance against antibiotics. Increasing antibiotic resistance amongst staphylococci makes it extremely challenging to treat these infections. Therefore, infections caused by staphylococci, especially *S. epidermidis* represent a serious burden on the health care system and hence it is important to characterize the virulence factors that are involved in these types of infections. The genes responsible for antibiotic resistance and novel virulence factors are acquired

through horizontal gene transfer (HGT). *S. epidermidis* is evolving through the addition of new virulence factors and antibiotic resistance genes. Increased knowledge of the mechanism of these novel virulence factors and their impact on establishing infections is the key to developing drugs to treat *S. epidermidis* infections.

S. epidermidis and other CoNS

General characteristics

Staphylococci are gram-positive bacteria, which under the microscope form round, grape-like clusters. There are 47 species and 23 subspecies under the genus *Staphylococcus*. The genus is divided mostly into two groups based on the presence of coagulase, a major virulence factor that causes clotting of blood: coagulase positive staphylococci (*S. aureus*, *S. pseudintermedius* and few others) and coagulase negative staphylococci (*S. epidermidis*, *S. haemolyticus* etc.). In addition, some staphylococci species are coagulase variable e.g. *S. schleiferi* includes both a coagulase positive subspecies (*S. schleiferi* subsp. *coagulans*) and a coagulase negative subspecies (*S. schleiferi* subsp. *scheiferi*) [2].

CoNS is a part of the commensal skin flora. *S. epidermidis* is the most frequently isolated staphylococcal species from humans, particularly from moist areas such as axillae, and anterior nares, toe webs and perineal areas [3, 4]. *S. capitis* is present around the forehead sebaceous glands [5]. *S. haemolyticus* colonizes axillae and pubic areas. *S.*

saprophyticus is isolated from the rectum and the genitourinary tract [4, 6]. CoNS was previously considered as a contaminant of blood cultures isolated from patients because of their ubiquitous presence on human skin.

Clinical significance

Presently, CoNS are regarded as opportunistic pathogens and are a major cause of nosocomial infections. *S. epidermidis* and *S. haemolyticus* are most frequently associated with diseases in humans compared to other CoNS. *S. epidermidis* is "medium" pathogenic staphylococci i.e. it expresses less virulent factors compared to *S. aureus*. However, *S. epidermidis* causes a different disease spectrum than *S. aureus*[2]. The pathogenic potential of *S. epidermidis* lies in its ability to colonize and infect every biomaterial developed to date. The most common isolates recovered from nosocomial central-line associated bloodstream infections (CLABSI), as recorded by National Healthcare Safety Network at the Centers for Disease Control and Prevention, were CoNS, with *S. epidermidis* responsible for half of these infections [7]. Studies have reported that patients with CLABSI had significantly longer ICU and hospital length of stay, higher mortality and hospital costs than uninfected patients [8, 9]. The estimated attributable cost per infection is $ 11,971 (2006 dollars) to $45, 814 (2012 dollars) [10, 11]. Other FBR-blood stream infections include infections associated with prosthetic heart valves and vascular grafts, cardiac devices, and coronary stents [2].

Other infections caused by *S. epidermidis* include native valve endocarditis, infections in

preterm infants and in very low birth infants, and bacteremia/septicemia in patients. CoNS is an emerging important cause of native valve endocarditis in both health care and community settings. In one study, data published from the International Collaboration on Endocarditis Prospective Cohort study involving 1635 patients from 61 centers in 28 countries, showed that CoNS was responsible for ~8% of these cases in patients with no history of injection drug use. It was also reported that native valve endocarditis due to CoNS is associated with poor outcomes, i.e. prolonged symptom duration and congestive heart failure. *S. epidermidis* alone caused 80% of the infections reported in this study [12]. In addition, CoNS frequently causes late-onset sepsis in neonates, especially very low birth weight infants. Amongst CoNS, *S. epidermidis* is the most prevalent pathogen in neonates [13, 14]. Furthermore, blood stream infection is a severe complication in patients with chemotherapy-induced neutropenia. CoNS was responsible for 40% of infections caused in a multicenter study of 1,051 bacteremia episodes in 782 cancer patients [15].

CoNS, other than *S. epidermidis* are also regarded as opportunistic pathogens. After *S. epidermidis*, amongst CoNS *S. haemolyticus* is the second most isolated pathogen from human blood cultures [16, 17]. *S. haemolyticus* has adapted very well to the hospital environment because of its genome plasticity and acquisition of antibiotic resistant determinants [18]. Additionally, *S. saprohyticus* is the second leading cause of urinary tract infection in young, sexually active women, after *E. coli* [19]. *S. capitis* is also emerging as important opportunistic pathogen in neonatal intensive care units. *S.*

capitis has been reported to cause more infections than *S. epidermidis*, in infants admitted in neonatal intensive care units [20].

In conclusion, an increase in the use of indwelling or implanted foreign bodies, and an increase in the number of patients with impaired immune system has established CoNS as one of the major nosocomial pathogens.

Biofilm

Biofilm is defined as a surface attached aggregate of bacteria forming a distinct architecture. Biofilm consists of bacterial communities embedded in a matrix where bacteria have unique transcriptional responses from those growing in the planktonic phase [21, 22]. Bacteria in staphylococci biofilm have heterogeneous states (aerobically and anaerobically growing cells, dormant and dead cells) due to spatial and temporal responses to their surrounding environment [23, 24]. This heterogeneity of biofilms is responsible for enhanced tolerance to antimicrobials like antibiotics, and antibodies [23]. Staphylococci have established themselves as successful nosocomial pathogens because of their ability to form biofilm on biomaterials. The process of forming a biofilm is divided into three stages: attachment, accumulation and dispersion (Figure 1).

The first step in biofilm formation is the attachment of the bacteria to abiotic or conditioned surfaces. This step is crucial for establishing infection. The expression of autolysin E (AtlE) and Autolysin/adhesion of *S. epidermidis* (Aae) changes cell surface

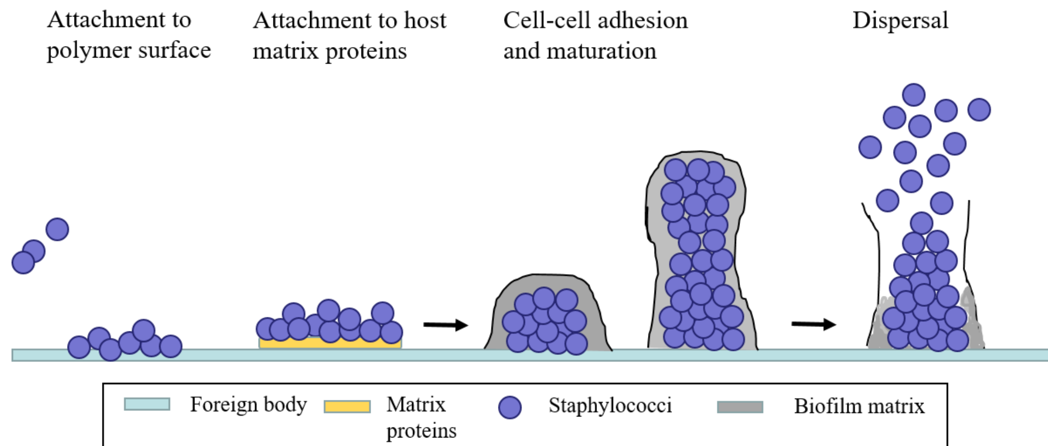


Figure 1. **Schematic representation of biofilm formation**. Adapted with permission from [25].

hydrophobicity, which mediates the attachment to abiotic surface through nonspecific and hydrophobic interactions [26, 27]. Once inserted in a host, human serum proteins e.g. fibrinogen, collagen and fibronectin quickly adhere to the biomaterial [28]. Not surprisingly, *S. epidermidis* expresses surface proteins that can mediate interactions with human serum proteins. Serine aspartate dipeptide repeat protein G (SdrG) and F (SdrF) belong to microbial surface components recognizing adhesive matrix molecules (MSCRAMM) family of surface proteins [29]. SdrG binds to human fibrinogen and this interaction has been shown to play a role in colonizing biomaterial [30, 31]. On the other hand, SdrF binds to α1 and α2 chains of Type 1 collagen [32]. Reports show that SdrF binds to unmodified Dacron surfaces covering drivelines as well as collagen coated on cardiac assist device drivelines [33]. In addition extracellular matrix binding protein (Embp) binds to fibronectin and potentially contributes to the attachment phase [34, 35]. Two bacterial autolysins also play a role in attachment. The autolysins AltE and Aae release eDNA into the matrix after cell lysis [36]. Extracellular DNA (eDNA) plays a role in primary attachment as addition of DNase I reduced bacterial attachment to glass surface [37].

After attachment, the next step of biofilm formation is known as accumulation, where intercellular adhesion molecules cause bacteria to aggregate and develop a multicellular, multilayered biofilm architecture [23]. Two critical factors shown to play a role in the accumulation phase are polysaccharide intercellular adhesion (PIA) and Accumulation associated protein (Aap). PIA is a homoglycan composed of β-1,6-linked 2-amino-2-

deoxy-2-glyucopyranol residues [38]. PIA is synthesized by the genes in the *ica* operon [39]. PIA plays a role in initial attachment, accumulation, and the architecture of the mature biofilm [23, 32]. Biofilm generated by PIA expressing *S. epidermidis* isolates have significant tower formation and 3D structure [23]. Although PIA is critical for biofilm formation, *S. epidermidis* isolates that lack the *ica* operon can still cause clinically significant infections. PIA independent biofilms are protein dependent [40]. The Aap protein has been shown to play a significant role in PIA independent or protein dependent biofilms. Aap is a cell wall anchored protein on the surface of *S. epidermidis* [41]. The B domains of Aap dimerize with B domains of Aap on the neighboring cells in a Zn^{2+} dependent manner causing cells to aggregate [42]. The Aap-dependent biofilms show less tower formation than the PIA dependent biofilms. The biofilm associated protein homolog (Bhp) and embp, cell wall anchored proteins in *S. epidermidis* also play a role in biofilm accumulation phase in absence of both PIA and Aap [35, 43].

The last stage of biofilm formation is known as dispersal. During dispersal, individual cells or sections of biofilm separate from the mature biofilm and migrate to other organs [44]. Phenol soluble modulin (PSM) are pro-inflammatory peptides regulated by the *agr* gene [45]. PSM-β inhibit noncovalent interactions between bacteria and the biofilm by acting as a surfactant and thereby promoting biofilm detachment [46]. Wang et al. [47]showed that *in vivo* PSM-β helps in dissemination of *S. epidermidis* from infected catheters to other organs. Agr, a two-component system, also regulates biofilm dispersal by upregulating proteases that cleave proteins involved in biofilm accumulation [48].

Virulence factors

Virulence factors are defined as factors that enable replication and dissemination of bacteria in the host by evading host immune response. Virulence factors can be divided in to different categories: (i) colonizing factors – molecules that allow bacteria to colonize host, (ii) adhesion factors – molecules that enable bacteria to adhere to host cells, (iii) invasion factors – molecules that allow bacteria to invade host cells, and (iv) evasion factors – molecules that allow bacteria to evade host immune responses by either subverting responses or killing host cells. *S. epidermidis* has lower virulence potential compared to *S. aureus* because it lacks the toxins encoded in *S. aureus* [24]. Instead, *S. epidermidis* relies on immune evasion factors, which promote persistence. *S. epidermidis* evades host response through biofilm formation, production of protective exopolymers, and sensing of antimicrobial peptides [24, 32]. Table1 is a summary of virulence factors encoded by *S. epidermidis* and their function.

Antibiotic resistance

Antibiotic resistance is widespread amongst clinical *S. epidermidis* isolates. CoNS appear to serve as a reservoir for genes encoding virulence factors that can be transferred horizontally to other staphylococcal species including *S. aureus*. Resistance to methicillin, rifamycin, fluoroquinolones, gentamicin, tetracycline, chloramphenicol, erythromycin, clindamycin and sulfonamides has been reported in *S. epidermidis*. 73-80% of *S. epidermidis* isolates obtained in the clinic are resistant to methicillin [49, 50].

Virulence factor	Function	Reference
AltE	Autolysin, binds to vitronectin	[26]
Aae	Autolysin, binds to fibrinogen, fibronectin and vitronectin	[27]
SdrF	Binds to collagen, initiation of ventricular assist device driveline infections	[31, 51]
SdrG	Binds to fibrinogen, attachment to catheters	[29, 30]
Embp	Binds to fibronectin, intercellular protein adhesin	[34, 35]
Ebp	Binds to elastin	[52, 53]
Bhp	Intercellular protein adhesion	[43]
PIA	Intercellular polysaccharide adhesion, immune evasion	[54, 55]
Aap	Binds to corneocytes, initial attachment in biofilm, intercellular protein adhesin	[56, 57]
Teichoic acids	Component of biofilm matrix	[58]
PGA	Forms outer capsule, protects from antimicrobial peptides and phagocytosis	[59]
SepA	Protease, antimicrobial peptide degradation	[60]
Antimicrobial peptide sensing system	Antimicrobial peptide sensor, regulates antimicrobial resistance mechanism	[61, 62]
Phenol soluble modulin (PSM) α	Cytolysin, pro-inflammatory	[63, 64]
PSM β1, PSM β2	Biofilm detachment, pro-inflammatory	[47]
PSM δ	Cytolysin, pro-inflammatory	[65, 66]
PSM ε	Cytolysin, pro-inflammatory	[65, 66]
δ-toxin	Cytolysin, pro-inflammatory	[65, 66]
PSM-mec	Cytolysin, pro-inflammatory	[65, 66]
SEC3, SEIL	Enterotoxins located in pathogenicity island	[67]
Fatty acid modifying enzyme	Detoxification of bactericidal fatty acids	[68]
SesC	Biofilm formation	[69, 70]

Table 1. Virulence factors of *S. epidermidis*.

MSCRAMMs

Cell wall-anchored (CWA) proteins of bacterial pathogens are important factors to establish adherence to host tissues, a critical step in the pathogenesis of microbial infections [28]. Bacterial CWA proteins are described as mosaic proteins composed of interlinked domains. These CWA proteins are grouped into families based on structural or functional homology [71]. Many CWA proteins of Gram-positive bacteria contain two linked IgG-like folded domains, which mediate ligand binding using the dock, lock and latch (DLL) model. These surface proteins are grouped into the Microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) family [29, 72].

All MSCRAMMs have a N-terminus signal sequence, an A-domain containing usually the ligand binding region, a C-terminus wall spanning region, LPXTG motif, a hydrophobic membrane spanning domain and a cytoplasmic tail. N-terminus signal sequences (~40 aa) is required for secretion through Sec-dependent pathway. The A-domain of staphylococcal MSCRAMMs is further divided into three sub-domains: N1, N2 and N3. The N2 and N3 sub-domains adopt IgG-like folds and represent the principal ligand-binding sub region. N2N3 sub- domains of MSCRAMMs bind ligand through the DLL mechanism. The cell wall spanning region is rich in either glycine and proline residues or serine and aspartate residues [28, 52]. MSCRAMMs are anchored to the cell wall through sortase enzyme, which recognizes the LPXTG motif present in surface

proteins. The sortase enzyme cleaves the bond between threonine and glycine residues. The carboxyl group of threonine is then covalently linked to the branch peptide in the peptidoglycan, anchoring the protein to the bacterial surface [73, 74]. Due to processing of the nascent peptide, the N-terminus signal sequence, the membrane spanning region and the cytoplasmic tail are not present in mature protein on the surface of the bacteria [74, 75].

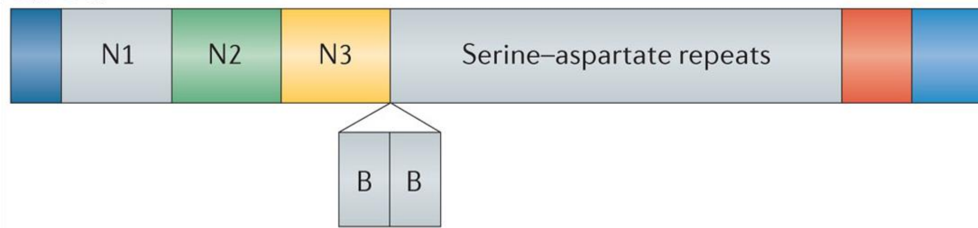
Three known families of MSCRAMM

Clumping factor-Serine aspartate dipeptide repeat (Clf-Sdr) family

In *S. epidermidis*, this sub-family of proteins contain SdrF and SdrG [76]. In *S. aureus*, ClfA, ClfB, SdrC, SdrD, SdrE, Bbp constitute Clf-Sdr sub-family [77]. Figure 2a shows relative positions of the N-terminus signal peptide, the A-domain (A), the B-repeat domains (B_n), the serine-aspartate dipeptide repeat region, the LPXTG motif and the cell wall/membrane- spanning region. Serine-aspartate dipeptide region, a characteristic feature of Clf-Sdr sub-family acts as linker to extend the A-domain away from the bacterial cell surface [78]. The SdrC/D/E/G/F and Bbp proteins contain B-repeats (~110 aa) [79]. With the exception of SdrF, these proteins bind to their ligands through N2N3 sub-domains of the A-domain [77]. In SdrF, the B-repeats instead of the A-domain bind to Type IV collagen [31]. The function of SdrF A-domain in unknown yet.

Fibronectin binding proteins (FnBPs)

FnBPA and FnBPB constitute the FnBP subfamily of MSCRAMMs in *S. aureus* [77].



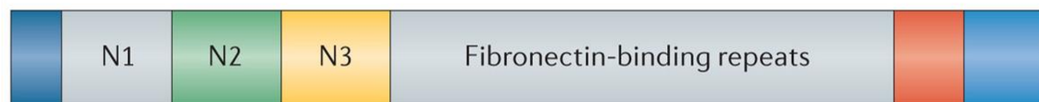


Figure 2. **Cartoon representation of previously characterized subfamilies of MSCRAMM.** Reprinted with permission from [29].

The cartoon in figure 2b shows relative positions of N-terminus signal peptide, A-domain (A), Fibronectin-binding repeats, LPXTG motif and cell wall/membrane-spanning region. As the name states, proteins in this subfamily bind fibronectin [80, 81]. Fibronectin-binding repeats in FnBPA and FnBPB act as a linker and function as a ligand binding region [28]. The N2N3 sub-domains of both proteins binds to fibrinogen [82]. The A-domain of FnBPs is structurally and functionally similar to Clf-Sdr proteins [83, 84].

Collagen adhesin (Cna) protein

The Cna protein is the only known *S. aureus* MSCRAMM in this subfamily [28]. The cartoon in figure 2C shows relative positions of N-terminus signal peptide, the A-domain (A), the B-repeats, the LPXTG motif and the cell wall/membrane- spanning region. Unlike other MSCRAMM, the N1N2 sub-domains of Cna comprises of a variant of the IgG fold. Cna binds to collagen by a collagen hug mechanism, which is a derivative of DLL mechanism [85]. Cna B-repeats differ in sequence from the Sdr proteins B-repeats [86]. Cna lacks a flexible linker to project ligand binding region away from cell surface.

MSCRAMM ligand-binding mechanisms

X-ray crystallography of recombinant N2N3 sub-domains with and without ligand, site directed mutagenesis, and amino acid replacements in the ligand peptide helped define the MSCRAMM ligand-binding mechanisms. There are two known mechanisms: (i) the dock, lock and latch model and (ii) the collagen hug model.

Dock lock and latch model

The N2 and N3 sub-domains of MSCRAMM each adopt an IgG-like fold, which is composed of two β-sheets. The N2 sub-domain contains A, B, E, D, D', D'', C, F, and G strands. The N3 sub-domain contains A', B', E', D', D1', D2', C', F' and G' [87]. The amino acid residues connecting the two IgG-like fold form a trench. Ligands dock into a trench formed between the two IgG-like folded domains. The ligand peptide aligns either parallel or antiparallel to G' strand in the N3 domain [28, 87, 88]. This docking induces a conformational change in the MSCRAMM resulting in a redirection of the C-terminal extension of the N3 domain to cover the bound peptide and locking it in the trench. Next, the N3 extension contacts the N2 domain by forming a complementary strand in a β sheet, thus forming a latch and a stable closed conformation of the ligand/MSCRAMM complex (Figure 3) [87, 88].

Collagen Hug model

Cna binds collagen using a collagen hug model, a variant of the DLL model. Collagen hug model consists of ligand docking, locking and latching as seen in DLL model. N1N2 sub-domains adopt IgG-like folds similar to N2N3 domains in other MSCRAMMs. Linker connecting N1 and N2 domain of Cna is longer than the linker between N2-N3 domains of other MSCRAMMs. Ligand docks in the cavity between the N1-N2 domain which results in a redirection of the linker to cover "hug" the bound peptide and locking it in place. Next, the C-terminal N2 extension contacts the N1 domain by forming a

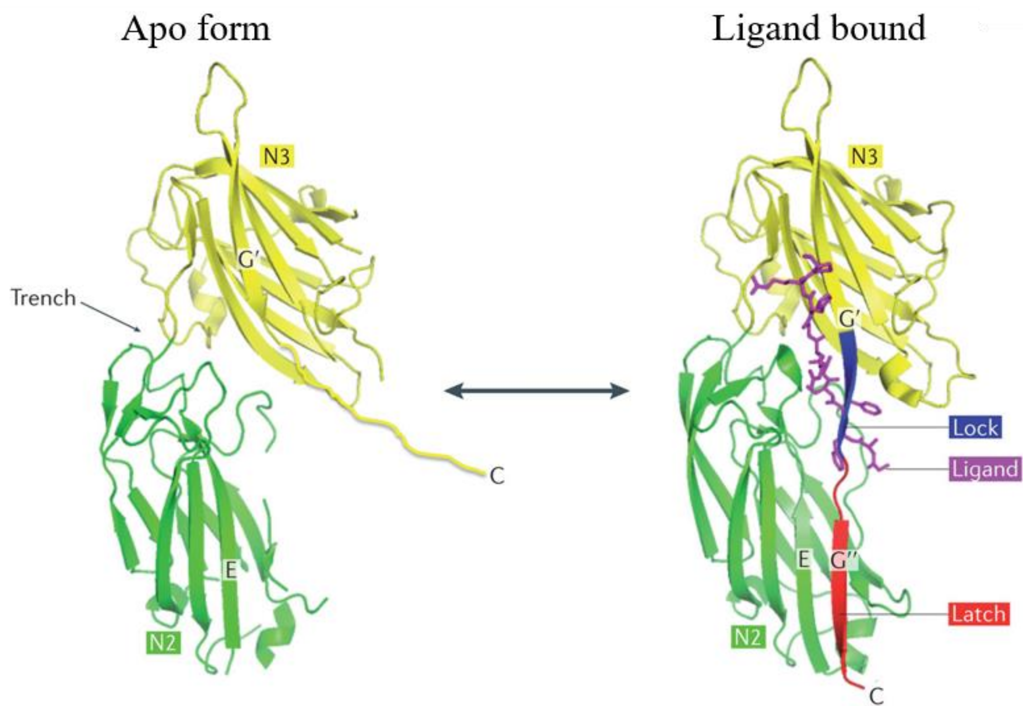


Figure 3. **Dock, lock and latch model.** Reprinted with permission from [29].

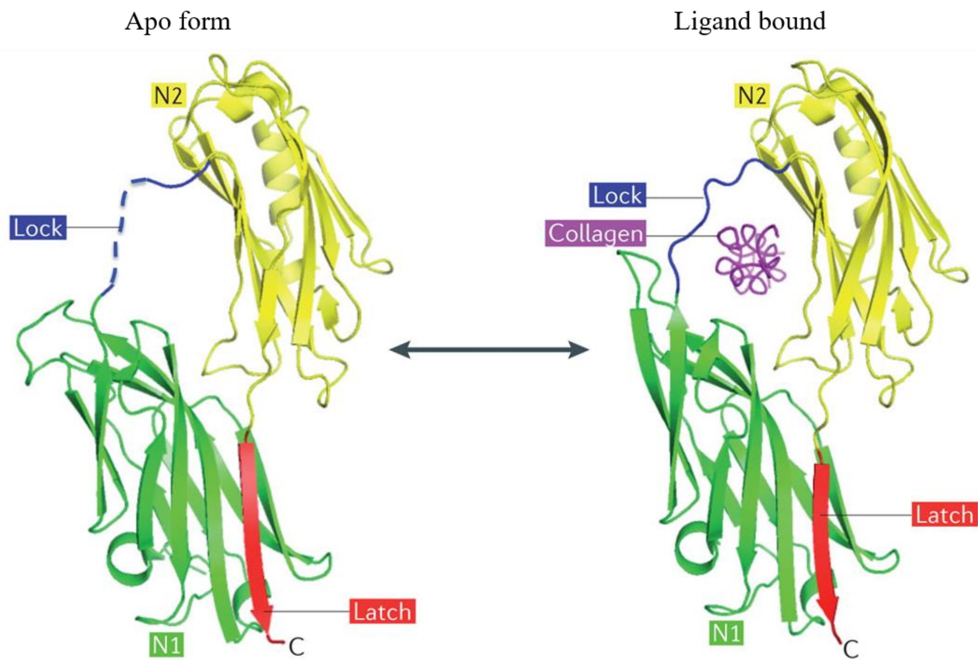


Figure 4. **Collagen hug model.** Reprinted with permission from [29].

complementary strand in a β sheet, thus forming a latch and a stable closed conformation of the ligand/MSCRAMM complex (Figure 4) [28, 85].

Function

Extensive studies of *S. aureus* MSCRAMMs has shown multiple functions for these surface proteins i.e. adhesion, invasion, immune evasion and biofilm formation. Table 2 provides a comprehensive list of MSCRAMM ligands and functions. 20% of the human population is colonized by *S. aureus* in the anterior nasal cavity. *S. epidermidis* is a ubiquitous skin colonizer [2]. MSCRAMMs bind to host proteins to accomplish two goals (i) colonization and (ii) adherence to host tissue to establish infections. ClfB binds to human loricrin expressed by squamous epithelial cells. This interaction promotes nasal colonization of *S. aureus* in the nose [89]. ClfB and ClfA both play a role in endocarditis, through their interaction with fibrinogen present in blood clot at a wound site [90, 91].

Recently, it has been discovered *S. aureus* can survive within neutrophils and even direct invasion of non-phagocytic cells. For example, FnBPA and FnBPB play a critical role in causing mastitis by promoting adhesion to epithelial cells in mammary gland and subsequent internalization [80, 81, 92]. Invasion of mammalian cells protect the bacteria from extracellular host defenses. In the case of *S. aureus*, internalization also provides an opportunity for bacterial toxins to damage the host cell from within. Upon encounter with a pathogen, the host launches innate and adaptive immune responses. To date,

MSCRAMM	Ligand/mechanism	Role in colonization or infection	Reference
ClfA	Binds fibrinogen in blood clot	Endocarditis	[91]
	Binds soluble fibrinogen	Immune evasion by coating bacteria with fibrinogen, septic death	[93, 94]
	Binds complement regulator Factor I	Reduces opsonophagocytosis	[95]
	Unknown mechanism	Critical for Septic arthritis	[96]
ClfB	Binds fibrinogen in blood clot	Endocarditis	[90]
	Binds to loricin expressed by squamous epithelial cells	Nasal colonization	[89]
	Cytokeratin 10	Nasal colonization	[97, 98]
FnBPA	Binds fibrinogen in blood clot	Endocarditis	[82]
	Binds to fibronectin	Invasion of epithelial cells in mammary gland, invasion of endothelial cells, mastitis	[80, 81, 92, 99]
	Binds to elastin	Adhesion to intra-aortic patch	[100]
	Homophilic interactions	Promotes biofilm on a biomaterial	[33, 101]
FnBPB	Binds to fibronectin	Invasion of epithelial cells in mammary gland, invasion of endothelial cells, mastitis	[80, 81, 92, 99]
	Homophilic interactions	Promotes biofilm on a biomaterial	[33, 101]
Cna	Collagen triple helix	Adhesion to cartilage tissue	[85, 102]
	Complement protein C1q	Inhibition of classical complement pathway	[103]
SdrC	β-neurexin	Unknown	[104]
	Homophilic interaction	Biofilm formation	[105]

Table 2. Functions of known MSCRAMMs.

MSCRAMM	Ligand/mechanism	Role in colonization or infection	Reference
SdrC	Adhere to desquamated epithelial cells	Unknown but points towards nasal colonization	[106]
SdrD	Adhere to desquamated epithelial cells	Unknown but points towards nasal colonization	[106]
SdrE	Binds to complement regulator factor H	Inhibit alternative complement pathway	[107]
Bbp	Binds to soluble and immobilized fibrinogen	Delays fibrin formation	[108]
SdrG	Binds to fibrinogen	Attachment to catheters	[29, 30]
SdrF	Binds to collagen	Initiation of ventricular assist device driveline infections	[31, 53]

Table 2. Continued

MSCRAMMs have been reported to interfere with innate immune responses. Cna binds to collagen like stalk of C1q and interferes with its interaction with C1r; thereby, inhibiting classical complement pathway [103]. Similarly, SdrE binds and recruits factor H to the surface of *S. aureus* and inhibits the alternative pathway of complement activation [107]. An alternative strategy *S. aureus* utilizes to evade immune response is by coating itself with fibrinogen. ClfA, ClfB, FnbpA, FnbpB, Bbp all bind to fibrinogen and can potentially create a fibrinogen coat around the bacteria, which shields it from innate and adaptive immune responses [82, 93, 108].

Lastly, MSCRAMM participate in biofilm formation through either hemophilic interactions or binding to a host protein. *S. epidermidis* MSCRAMM SdrG binds to both soluble fibrinogen and adsorbed fibrinogen on biomaterial, thus promoting biofilm formation [29, 30]. Additionally, *S. epidermidis* MSCRAMM SdrF initiates biofilm formation by binding to collagen adsorbed on left ventricular assist driveline [31, 33]. On the other hand, SdrC, FnbpA promote biofilm formation through homophilic interaction [101, 105]. N2-N3 sub-domain of FnbpA interacts with N2-N3 sub-domain of another molecule of FnBPA in a DLL independent mechanism [33, 101].

Mobile genetic elements

The *S. epidermidis* genome consists of 80% core genes and 20% variable genes [109]. The core genes are the genes present in all the strains of a bacterial species. Core

genome consists of genes required for survival e.g. DNA synthesis, replication etc. Variable genes are present in a single or a subset of strains of a bacterial species. Variable genes are often encoded on mobile genetic elements (MGEs) and include transcriptional regulators, defense mechanism genes, recombinase and integrase genes [109]. MGEs are segments of DNA with intracellular (within the same genome) and intercellular (between bacterial cells) mobility [110]. These elements encode proteins and enzymes that facilitate their transfer as well as integration into DNA of the new host cell. Bacteria adapt by acquiring MGEs encoding a variety of virulence and resistance determinants required for surviving selective pressures [110]. MGEs are propagated by vertical gene transfer (progeny cells inherit DNA from parent cells) or horizontal gene transfer (HGT) [111, 112]. Staphylococci contain types of MGEs described below.

Plasmids

These are circular or linear, double stranded, self-replicating DNA molecules. Plasmids are smaller than the host chromosome and do not contain genes required for the bacteria survival. Each *Staphylococcus* cell carries one or more plasmids with varied gene content [112]. Plasmids usually carry genes for their own replication and variable genes that are not encoded on host chromosome. In staphylococci, plasmids are either transferred through transduction or conjugation (discussed below). Upon entering a new host cell, plasmid can either remain as its own unit or integrate into host genome. Staphylococcal plasmids vary in size from 1.3 kb to 64.9 kb with majority of the plasmids in 20-30 kb size range.

Plasmids are notorious for carrying antibiotic resistance determinant genes. *S. aureus* recently acquired vancomycin resistance gene cluster from enterococci through transfer of Tn1546 encoded on a conjugative plasmid [113]. The pSTS7 plasmid from *S. epidermidis* mediates resistance to tetracycline, kanamycin and neomycin. pSTS7 has partial sequence homology to kanamycin/bleomycin/neomycin resistance plasmid pUB110 from *S. aureus* and tetracycline resistance plasmid pNS1981 from *Bacillus subtilis* [114]. Additionally, plasmid pNE131 from *S. epidermidis* encodes a methylase responsible for resistance to antibiotics macrolide, lincosamide and streptogramin B (MLS B) [115].

Bacteriophages

Bacteriophages, often-called phages, are viruses that infect bacteria. The phage genome consists of double stranded or single stranded DNA or RNA. Their size varies from a few to several 100 kb [110]. Phages core genome consists of genes required for forming head, tail, DNA replication and nucleotide metabolism. These genes are used to identify phage in a bacterial genome [116]. Phages can have lysogenic or lytic life cycles. Lytic phages replicate inside bacterial host and lyse the bacterial cell for the release of new viral progeny. Alternatively, in lysogenic life cycle, phage genome is maintained stably in the new host as plasmid or gets integrated in the genome (prophage). Phage DNA then replicates with the host genome. Under certain stress conditions, prophage can exit lysogenic cycle and turn lytic [117].

Prophages can also carry important virulence factors in their genome. Chemotaxis inhibitory protein (CHIP), staphylococcal inhibitor of complement (SCIN), leukocidin M/F, Pantone-Valentine leucocidin, staphylokinase, exfoliative toxin A, enterotoxin A/G/K/K2/P/Q are critical virulence factors encoded on prophage in *S. aureus* [112]. The ϕSPβ-like prophage in *S. epidermidis* encodes SesI, which has been suggested as a potential marker for invasive capacity of *S. epidermidis*. SasX, a homologue of SesI in *S. aureus* is also encoded on a similar ϕSPβ-like prophage [118]. The SasX protein in *S. aureus*, enhances nasal colonization, skin and lung disease, and abscess formation [119].

Transposons and insertion sequences

Transposons also called as “jumping genes”, are segments of DNA that move around in the genome. Transposons can be inserted in core genome or mobile genetic elements like plasmids and genomic islands. When inserted within a gene, the transposon can disrupt gene expression or function. Transposons can also pick up surrounding genes and move them to different genomic site.

Common transposons found in bacteria contain genes for antibiotic resistance. The Tn916/Tn1545 family consists of two related conjugative transposons discovered in the 1970s. Variations of Tn916/Tn1545 have been reported in multiple bacterial species including *S. aureus* and CoNS [120-122]. Tn1545 is a ~ 25 kb transposon carrying resistance determinants MLS B-type antibiotics and kanamycin. Tn916 is 18kb in size, and confers resistance to tetracycline and its analog minocycline [120]. Staphylococci

are notorious for carrying resistance to beta-lactam antibiotics. The β-lactamase structural gene and its regulatory genes have been identified on transposon Tn552 as well as plasmids and chromosome. Tn552 is closely related to other transposons like Tn4002, Tn4201 found in staphylococci [123]. Sidhu et. al. reported presence of Tn552 in *S. epidermidis* [124].

Genomic Islands

Genomic islands are clusters of genes acquired through horizontal gene transfer. Genomic islands share characteristic features: (i) Genomic islands can be anywhere from 8 kb to 200 kb. (ii) usually have a sequence composition bias i.e. GC content is different than GC content of the core genome. (iii) They are often inserted in *tRNA* genes (iv) Insertion of genomic islands creates 16–20 bp flanking direct repeats. (v) and frequently have a cluster of genes that offer a selective advantage for the bacteria [125, 126]. These genes could encode for virulence factors, fitness, resistance, metabolism or symbiosis. Certain genomic islands also encode for an enzyme required for its mobility like integrases, transposase and chromosomal cassette recombinases. In other cases, genomic islands could have remnant of mobility genes indicating these were once mobile. The staphylococci genome has multiple types of genomic islands: immobile genomic islands, pathogenicity islands, Arginine catabolic mobile element (ACME) and Staphylococcal chromosomal cassette (SCC) [112].

Sequencing and assembly of the complete genome of *S. epidermidis* helped to identify

υSe1, υSe2, a novel genomic island υSeγ and enterotoxin bearing pathogenicity island SePI [69, 127]. The υSe1 in *S. aureus* carries enterotoxin genes while in *S. epidermidis* RP62a it contains genes for cadmium resistance. *S. aureus* υSe2 encodes for SEC, TSST toxins. On the other hand, υSe2 in *S. epidermidis* ATTCC12228 isolate contains gene for *srtC*, strain specific sortase and two LPXTG motif containing surface proteins predicted to play a role in adhesion to host tissue. υSeγ encoded by both the *S. epidermidis* isolates contains gene cluster for four members of PSM family [127]. Recently, enterotoxin *sec3* and *sell* bearing pathogenicity island SePI was discovered in *S. epidermidis* FRI909. Out of 200 isolates tested, SePI was only present in one isolate, indicating its rare presence in *S. epidermidis* [69].

ACME is a genetic island first found in *S. epidermidis* ATCC12228 genome [55, 128]. The *arc* genes and the *opp* genes are the two main gene clusters present in ACME [128, 129]. The *arc* genes encode for enzymes of the arginine deiminase pathway which converts L-arginine to ATP, carbon dioxide and ammonia [128]. The *opp* genes encode for oligopeptide permease operon. The Opp operon has multiple functions including chemotaxis, quorum sensing, peptide nutrient uptake and more. It is worth noting that both *arc* and *opp* genes are encoded on the chromosome in staphylococci and other bacteria [130, 131].

In *S. aureus*, ACME has been identified in ST 5, ST22 and ST8 [128, 129, 132]. ACME island enhances pathogenic fitness (enhanced colonization and transmission) and is

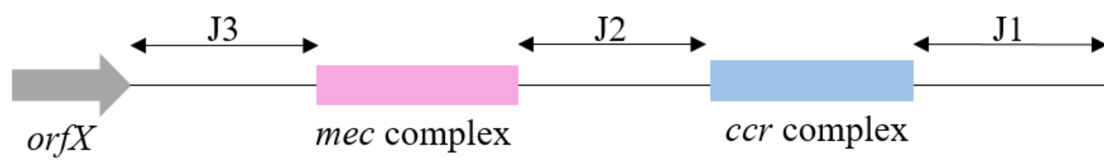


Figure 5. **Structural organization of SCC*mec*.**

responsible for USA300 *S. aureus* epidemic in the US [133, 134]. However, it is more widespread in CoNS especially *S. epidermidis*. ~50% of clinical isolates of *S. epidermidis* tested had ACME [135, 136]. ACME is often found together with SCC*mec* (discussed below) inserted either downstream or upstream of SCC*mec* [128, 137]. ACME is inserted at the *orfX*, in the same location as SCC*mec* [128, 136]. There are three allotypes of ACME: ACME I containing both the *opp* and *arc* clusters, ACME II containing the *arc* gene cluster and ACME III containing the *opp* gene cluster [136].

SCC is another type of genomic island present in staphylococci. There are two types of SCC: SCC*mec* (containing the *mec* gene) and Non-*mec* SCC. All SCC elements share conserved features: (i) always inserted in 3' of the *orfX* gene encoding rRNA methyltransferase, (ii) contain *ccr* genes encoding for site specific recombinases and (iii) are flanked by direct repeats. SCC*mec* contains genes for methicillin resistance complex (*mecA*, *mecR* and *mecI* genes) and chromosomal cassette recombinase complex (*ccrA/B* or *ccrC* genes). Methicillin resistance gene complex and chromosomal cassette recombinases genes are connected to each other and to the ends of the cassette by joining regions (J1, J2 and J3) as shown in figure 5. Joining regions encode for genes. Joining regions have also been shown to contain genes for resistance to antibiotics and heavy metals, and transposons [138]. SCC*mec* is typed based on the *mec* gene complex, *ccr* gene complex and sub-typed by J region. Table 3 lists the composition of *ccr* gene complex and *mec* gene complex present in 11 SCC*mec* identified to date. SCC*mec* types

SCC*mec* type	***ccr* gene complex**	***mec* gene complex**
I	1 (A1B1)	B (IS431-*mecA*-Δ*mecR1*-IS1272)
II	2 (A2B2)	A (IS431-*mecA*-*mecR1*-*mecI*)
III	3 (A3B3)	A (IS431-*mecA*-*mecR1*-*mecI*)
IV	2 (A2B2)	B (IS431-*mecA*-Δ*mecR1*-IS1272)
V	5 (C)	C2 (IS431-*mecA*-Δ*mecR1*-IS431)*
VI	4 (A4B4)	B (IS431-*mecA*-Δ*mecR1*-IS1272)
VII	5 (C)	C1 (IS431-*mecA*-Δ*mecR1*-IS431)#
VIII	4 (A4B4)	A (IS431-*mecA*-*mecR1*-*mecI*)
IX	1(A1B1)	C2 (IS431-*mecA*-Δ*mecR1*-IS431)*
X	7(A1B6)	C1 (IS431-*mecA*-Δ*mecR1*-IS431) #
XI	8(A1B3)	E (*blaZ*-*mecA*LGA251-*mecR1*LGA251-*mecI*LGA251)

* Two IS431s are arranged in the opposite direction

Two IS431s are arranged in the same direction

Table 3. The *ccr* and *mec* gene complex present in SCC*mec* types.

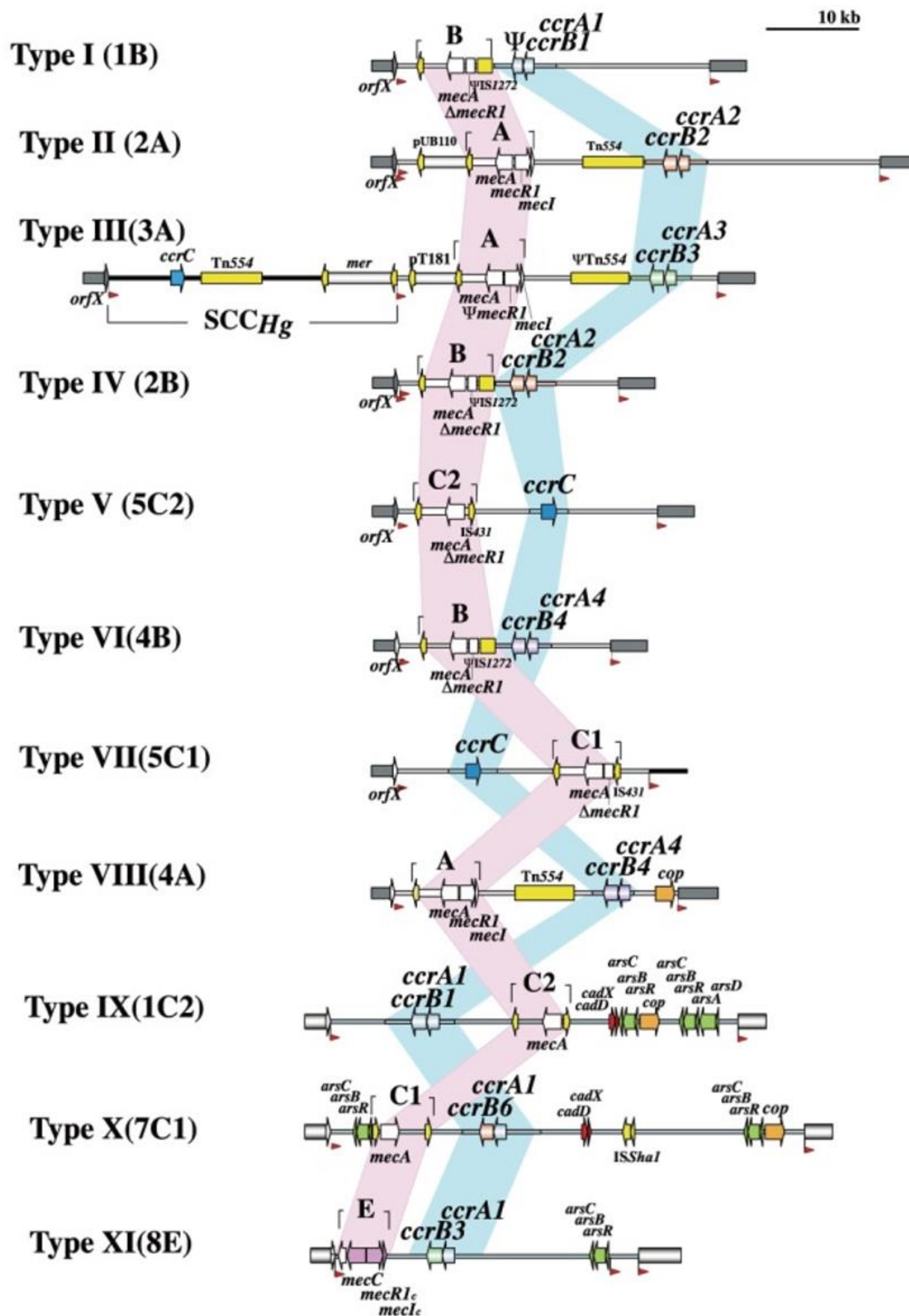


Figure 6. Schematic representation of SCC*mec* types. Reprinted with permission from [138].

I-VIII are present in isolates obtained from humans. Structures of SCC*mec* are shown in Figure 6 [138].

Peptidoglycan, a major component of the bacterial cell wall, is a polymer of long glycan chains connected by flexible peptide bridges. Glycan chains are made of repeating subunits of β-1,4-linked N-acetyl glucosamine (GlcNAc) and N-acetyl muramic acid (MurNAc) [139]. Peptidoglycan synthesis takes place in three different locations in a cell. First in the cytoplasm, Uridine diphosphate (UDP)-GlcNAc and UDP-MurNAc are generated. A five amino acid chain L-ala-D-Glu-L-Lys-D-ala-D-ala is added to UDP-MurNAc, which generates parks nucleotide. Second stage occurs at the plasma membrane where parks nucleotide is attached to a lipid anchor called undecaprenol pyrophosphate, bound to the plasma membrane. Parks nucleotide attached to lipid anchor is called Lipid I. Lipid I undergoes further modification by addition of N-acetyl glucosamine and a pentapeptide chain, creating Lipid II molecule. Glycine pentapeptide chain is added to L-Lys of Lipid I molecule. Next, lipid II molecule is flipped over to the outer side of the plasma membrane. Lipid II molecule is attached to the reducing end of an existing peptidoglycan chain via a transglycosylation reaction. Further, pentaglycine peptide on one Lipid II molecule is crosslinked with 4-D-Ala in the five amino acid chain of the other Lipid II molecule by a transpeptidation reaction. Transglycosylation and transpeptidation reactions are carried out by Penicillin binding protein (PBP) (Figure 7) [140].

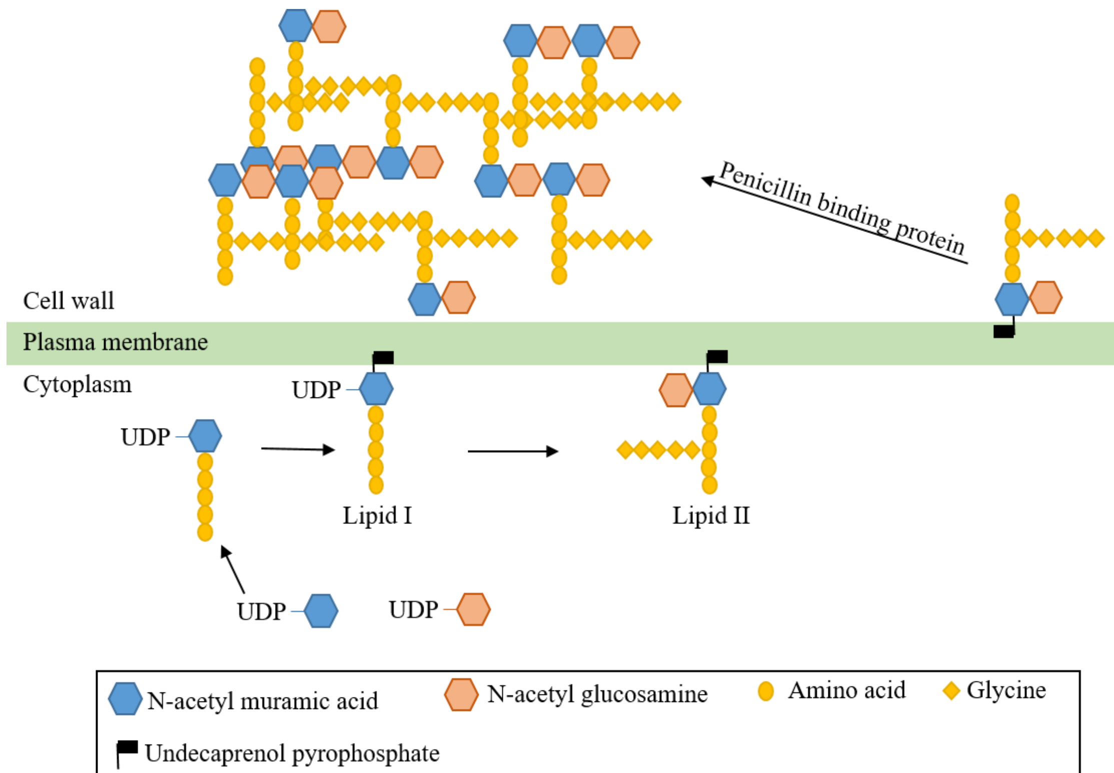


Figure 7. **Synthesis of peptidoglycan for cell wall formation.** Reprinted with permission from [140].

Methicillin is a β-lactam antibiotic of the penicillin class, which inhibits bacterial cell wall synthesis. Methicillin inhibits transpeptidation of glycan chains by acting as an analog of D-Ala-D-Ala molecules in the five amino acid chain in the Lipid II molecule [141]. A complex is formed between methicillin and PBP, which inhibits PBP function as a transpeptidase enzyme. Resistance to penicillin (another β-lactam antibiotic) is conferred by β-lactamase enzyme that cleaves a bond in the β-lactam ring of penicillin and deactivates the molecule. Methicillin escapes cleavage by by β-lactamase through steric hindrance and is therefore, able to target PBP [142].

A new Penicillin binding protein 2' (PBP2') confers resistance to methicillin. PBP2' lacks a high affinity site for binding to β-lactam ring of methicillin. In addition, it has a reduced rate of complex formation with methicillin compared to PBP [143]. PBP2' has an extended structure compared to PBP and is composed of transpeptidase domain with an allosteric site and a transmembrane domain. The binding of D-Ala-D-Ala terminus of pentapeptide stem provides the allosteric control. Binding of D-ala-D-ala to the allosteric site leads to conformational changes in the protein through formation of salt bridges. This conformational change opens up the active site for transpeptidation reaction [144].

The *mecA* gene in SCC*mec* encodes PBP2'. The *mecA* gene expression is increased in the presence of methicillin. The *mec* operon consists of four genes: *mecA*, *mecR1*, *mecI* and *mecR2* (Figure 8). The *mecA* gene is transcribed in opposite direction to the other genes in the operon. In the absence of methicillin, expression of the *mecA* gene is

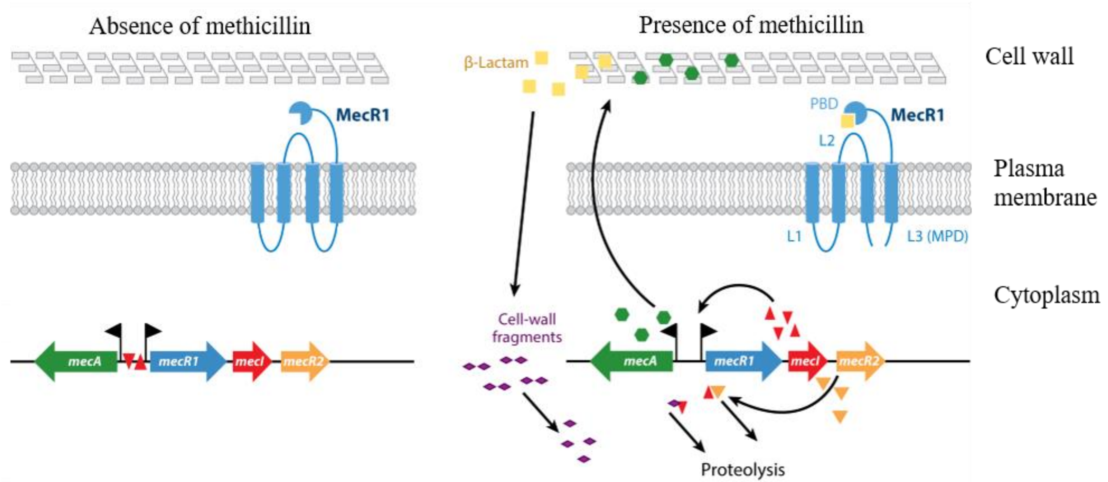


Figure 8. **The *mec* operon and its expression in absence/presence of methicillin.**

Reprinted with permission from [142].

inhibited by binding of the MecI repressor at *mec* operator site. MecR1 is a transmembrane protein that recognizes methicillin. Upon binding to methicillin, MecR1 undergoes autocatalytic activation and might be involved in proteolysis of MecI. D-Glu-D-Lys dipeptide cell wall fragments, generated by disruption of cell wall synthesis bind to MecI. This also leads to its dissociation from *mec* operator site, followed by transcription of the genes producing PBP2', MecI, MecR1 and MecR2. MecR2 is an anti-repressor that binds to MecI, leading to its proteolysis and prevents its binding to *mec* operator site [142].

SCC*mec* also encodes its own enzymes required for mobility. Once inside the host, chromosomal cassette recombinases (CcrA/B or CcrC) are required for integration and excision of the SCC elements. CcrA and CcrB complex is needed for effective transposition. CcrA is involved in substrate recognition and recombination while CcrB is involved in recombination only. Whereas, CcrC can both integrate and excise SCC elements by itself. *orfX* contains an integration site sequence attB at the 3' recognized by Ccr recombinases. Similarly, an attachment site called attSCC is present in circularized SCC*mec* element. Recombinases bring the two sites closer by formation of a tetrameric complex. Following DNA cleavage, 180° rotation of two of the Ccr molecules in the tetrameric complex leads to exchange of DNA strands. Re-ligation of DNA leads to integration of SCC*mec* element in the bacterial chromosomal DNA creating two new attachment sites called attL and attR. Excision of SCC*mec* is carried out in similar way and generates the original attachment sites attB and attSCC (Figure 9). Insertion of

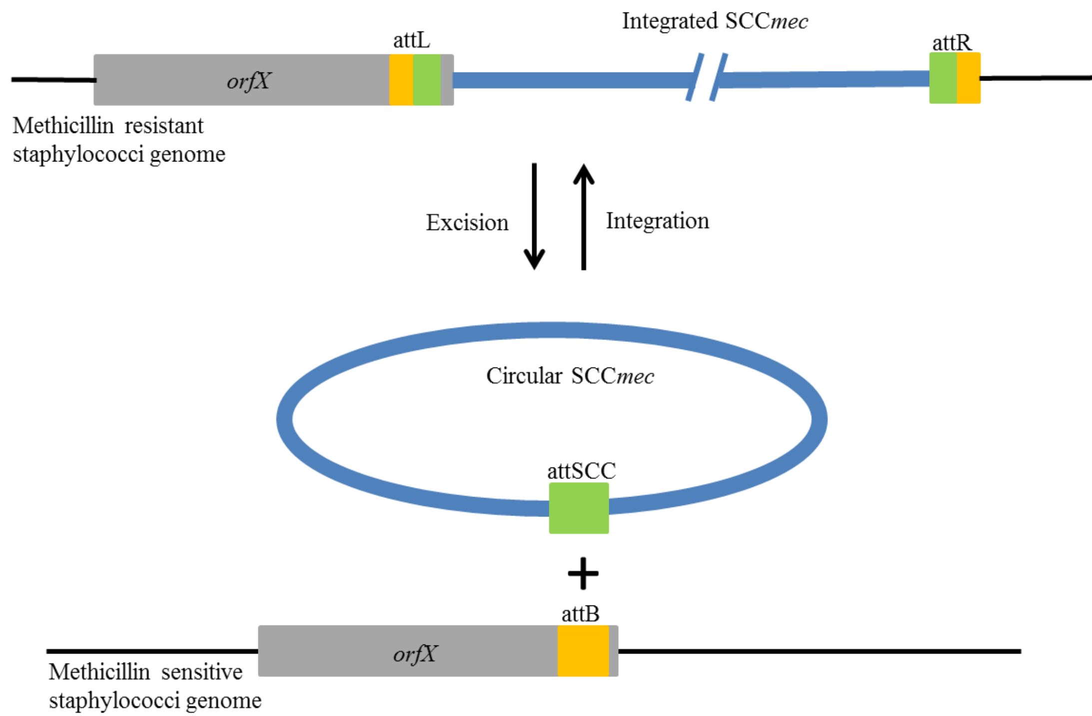


Figure 9. **Integration and excision of SCC*mec* elements.**

SCC*mec* cassette in *orfX* does not change its expression and the stop codon is not altered [145].

Horizontal gene transfer

MGEs are transferred amongst bacterial cells through vertical gene transfer or HGT. In prokaryotes, HGT occurs in 3 ways in prokaryotes: transformation, conjugation and transduction (Figure 10). Transformation is the direct uptake of exogenous genetic material from the surrounding. Staphylococci unlike *Bacillus subtilis* is not very competent [112, 117]. However, recently reported competence genes in *S. aureus*, are controlled by the expression of a novel secondary sigma factor, SigH. The report showed that a complete SCC*mec* type II can be transferred from N315 MRSA strain to MSSA strain through natural competence. SigH is only expressed in a fraction of bacterial cells [146].

Conjugation is the transfer of genetic material between bacterial cells by direct contact via pilus. Conjugation requires either conjugative plasmids or chromosomally integrated conjugative elements (ICEs) including transposons. Conjugative plasmids carry conjugative or transfer genes to establish mating pair and transfer plasmid from donor to recipient cells. Conjugation plasmids can also mobilize plasmids that contain limited or no conjugative genes. pSK41 is one of the most studied conjugative plasmid families in *S. aureus*. Conjugative pili have not been observed for *S. aureus* and it is therefore

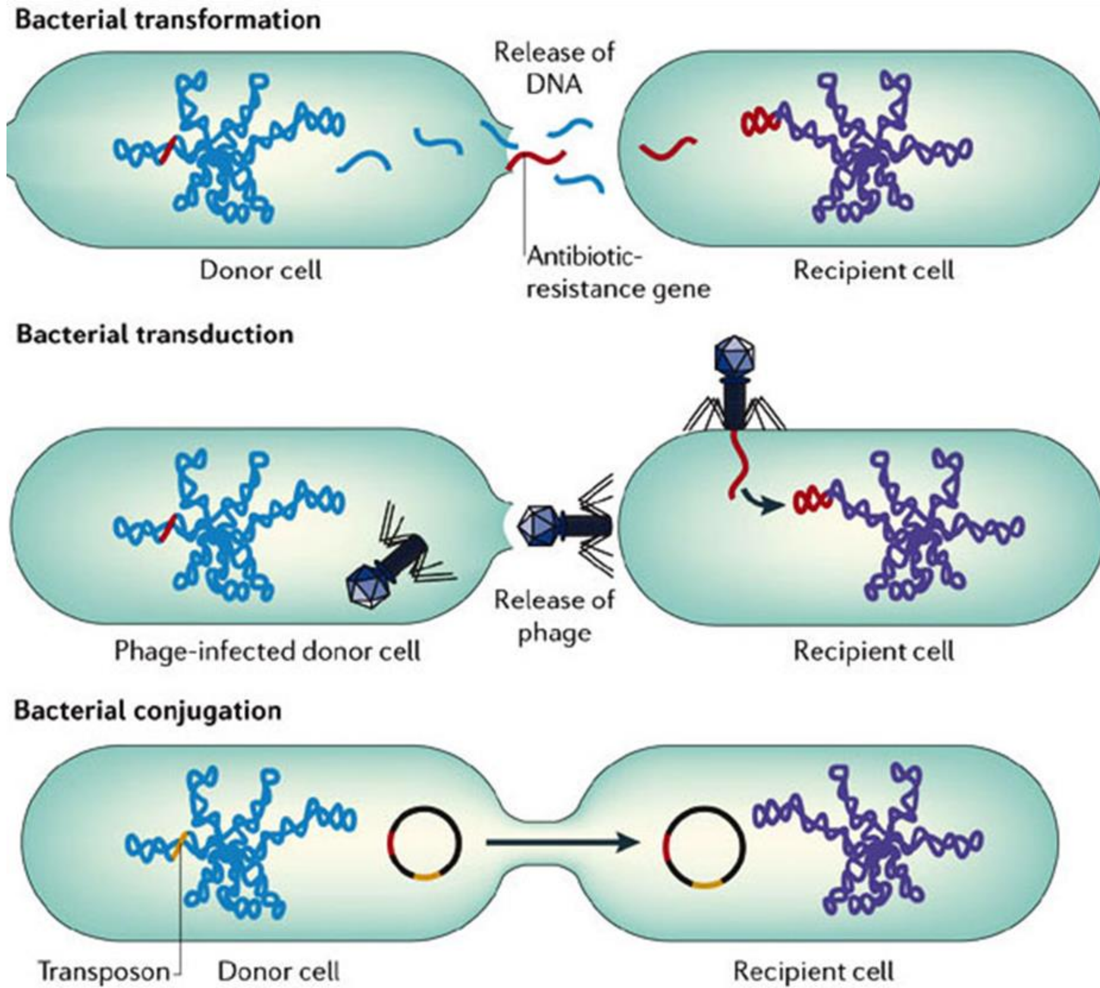


Figure 10. **Mechanisms of horizontal gene transfer in prokaryotes.** Reprinted with permission from [147].

believed that DNA is transferred through pores in neighboring bacterial cells. Also, frequency of conjugation is *S. aureus* is higher in biofilms compared to planktonic cells [11].

Transduction is the process of transfer of DNA from one bacterial cell to another by a bacteriophage. Phage typically packs its own DNA into the capsid, but under unknown conditions bacterial chromosomal or plasmid DNA can be packaged into the phage particle. This pseudo-phage particle when released from infected bacterial cell can bind to the receptor of an uninfected cell and transfer bacterial chromosomal or plasmid DNA. This mechanism is called generalized transduction. Newly inserted DNA lacks phage genes and is then integrated into the host bacterial cell genome and do not trigger a lytic lifecycle [148]. Scharn C.R. et. al. reported that bacteriophage 80 alpha and 29 can transfer SCC*mec* type IV and SCC*mec* type I respectively from USA 300 MRSA strains to methicillin-sensitive USA 300 recipient strains. This transduction occurred at the frequency from $3.8x10^{-9}$ to $6.9x\ 10^{-10}$. Transduction was dependent on presence of penicillinase (beta-lactamase specific for penicillin) plasmid in the recipient strain [149].

CoNS as a reservoir of antibiotic resistance

S. aureus is more virulent that CoNS because of plethora of virulence factors present in its genome. Most of the antibiotic resistance determinants and virulence factors in *S. aureus* are encoded on MGEs. CoNS appear to serve as a reservoir for genes encoding

virulence factors that can be transferred horizontally to other staphylococcal species including *S. aureus* [150]. Different staphylococcal species share ecological niches during their lifestyle as a commensal on human skin. This provides an opportunity for MGE exchange through HGT. Widespread *mecA* gene in *S. aureus* originated in *S. sciuri* [151].

There is strong evidence suggesting transfer of SCC*mec* elements from *S. epidermidis* to *S. aureus*. *S. epidermidis* SCC*mec* type IV and *S. aureus* SCC*mec* type IVa share a 98-99% homology at nucleotide level although their genomes do not share that high identity [152]. Insertion element IS1272 is present in higher numbers in *S. epidermidis* than *S. aureus* [153]. Additionally, SCC*mec* type IV was first observed in *S. epidermidis* [154]. Also, arginine catabolic mobile element (ACME) that contributes to fitness of Community associated-MRSA USA300 and a surface protein SasX located on a prophage in *S. aureus* originated from *S. epidermidis* [150].

CHAPTER II

A NOVEL MSCRAMM SUBFAMILY IN COAGULASE NEGATIVE STAPHYLOCOCCAL SPECIES[1]

CoNS are important opportunistic pathogens. *S. epidermidis*, a coagulase negative staphylococcus, is the third leading cause of nosocomial infections in the US. Surface proteins like MSCRAMMs are major virulence factors of pathogenic gram-positive bacteria. Here, we identified a new chimeric protein in *S. epidermidis*, that we call SesJ. SesJ represents a prototype of a new subfamily of MSCRAMMs. Structural predictions show that SesJ has structural features characteristic of a MSCRAMM along with a N-terminal repeat region and an aspartic acid containing C-terminal repeat region, features that have not been previously observed in staphylococcal MSCRAMMs but have been found in other surface proteins from gram positive bacteria. We identified and analyzed structural homologs of SesJ in three other CoNS. These homologs of SesJ have an identical structural organization but varying sequence identities within the domains. Using flow cytometry, we also show that SesJ is expressed constitutively on the surface of a representative *S. epidermidis* strain, from early exponential to stationary growth phase. Thus, SesJ is positioned to interact with protein targets in the environment and plays a role in *S. epidermidis* virulence.

---

[1] This chapter is reprinted with permission from "A Novel MSCRAMM Subfamily in Coagulase Negative Staphylococcal Species" by Arora, S., Unhlemann, A.C., Lowy, F.D., and Hook, M., 2016. *Frontiers in Microbiology*, **7**: 1-9, Copyright 2016 by Srishtee Arora, Ana-Catrine Uhlemann, Franklin D. Lowy and Magnus Hook.

Introduction

CoNS, which colonize human skin and mucus membranes, are recognized as important opportunistic pathogens. *S. epidermidis* and *S. haemolyticus* are the two most prevalent CoNS species responsible for causing a significant proportion of device-related, health care-associated infections and infections in preterm newborns [2]. *S. epidermidis* alone is the third leading cause of nosocomial infections in the US. *S. saprophyticus* is the second most common cause of uncomplicated urinary tract infections in sexually active women between the age of 18 and 35 years [19, 155]. Recent studies have indicated that *Staphylococcus capitis* can cause late onset sepsis in very low birth weight infants in the neonatal intensive care unit setting [156, 157] and prosthetic valve endocarditis in adults [158].

Bacterial surface proteins such as CWA proteins have been identified as important virulence factors among gram positive bacterial pathogens and play key roles in microbial adherence to host tissues, evasion of host defense systems and biofilm formation [29, 159]. For example, in *S. epidermidis* the CWA proteins Biofilm associated protein (Bap) [43], Aap [160], SdrF [53], SdrG [30] and Embp [161] all can participate in biofilm formation either by mediating bacterial attachment to matrix proteins or intercellular aggregation. Furthermore antibodies to the *S. epidermidis* surface protein C (SesC) inhibit biofilm formation, but a molecular function for SesC in biofilm formation has not yet been determined [72, 162].

The CWA proteins of gram-positive bacteria are often modular proteins composed of various interlinked domains. These proteins can be further divided into families [e.g., MSCRAMMs and Serine Rich Repeat Proteins (SRRPs)] based on the presence of common characteristic domains [29, 51]. A defining feature of the MSCRAMM family is the presence of two tandemly linked IgG-like folded domains, which can engage in ligand binding by the DLL mechanism [87, 88]. A subfamily of MSCRAMMs contains a repeat (R) region composed of serine-aspartate di-peptide repeats (Sdr), which defines the Sdr protein subfamily. In *S. aureus,* this subfamily includes ClfA, ClfB, SdrC, SdrD and SdrE [29, 77]. There are two members of the Sdr subfamily on the surface of *S. epidermidis*; SdrG and SdrF [76]. SdrG binds to human fibrinogen [87]. SdrF binds to Type I Collagen and is involved in the initiation of left ventricular assist device driveline infections [53].

The SRRP family of CWA proteins is defined by the presence of a serine repeat region (SRR), which is similar to the R region of Sdr proteins in that it contains serine dipeptide repeats but with either alanine, valine or threonine as the partner residue. SRRPs have a N-terminal signal sequence, at most two unique non-repeat (NR) regions, two SRRs flanking the NR region(s) and motifs needed for cell wall anchoring at the C-terminus [51]. Based on the crystal structures of NRs of two SRRPs; Fap1 from *S. parasanguinis* and GspB from *S. gordonii*, the NR region can be further subdivided into different domains. Both proteins contain one IgG-like folded sub-domain in the NR region [163, 164]. The SRRPs are common in streptococci but also have been found in *S. aureus*

(SraP in strain N315, [165]), *S. haemolyticus* (SH0326 in strain JCSC1435, [18]) and *S. epidermidis* (SE2249 in strain ATCC 12228, [166]).

Although the pathogenic mechanisms of *S. epidermidis* are attracting more attention, the possible roles of CWA proteins in these infections need to be further examined. Furthermore, CWA proteins on other CoNS are even less well characterized. We here report on the discovery of a previously unknown CWA *S. epidermidis* protein that we demonstrate is a prototype of a novel subfamily of CoNS MSCRAMMs.

Materials and methods

Bacterial strains and growth conditions

S. epidermidis strains were routinely grown in Tryptic Soy Broth (TSB) medium overnight at 37°C at 200 rpm. Growth curves were generated by inoculating fresh TSB media with overnight inoculum to a starting OD600 of 0.03, followed by incubation at 37°C at 200 rpm. *S. epidermidis* strain 3094 and 2111 are clinical isolates obtained from patients with left ventricular device driveline infections. *S. epidermidis* 3094 was used as a source for cloned constructs. *Escherichia coli* strains XL1Blue and BL21 AcellaTM were grown in Luria-Bertani (Sigma) medium with appropriate antibiotics at 37°C.

Identification of SesJ structural homologs

Bacterial genome sequences in the NCBI database were searched for proteins with

similarity to the SesJ protein using BLAST. New proteins were only accepted if they contained N-terminal Repeats (NTRs) at the N-terminus and repeats at the C-terminus. Search results were further analyzed to select for proteins that contained a N-terminus signal sequence, A-region, B repeats, a LPXTG motif, transmembrane domain followed by a positively charged C-terminal amino acid sequence. Online bioinformatics tools were used to characterize protein sequences from the BLAST search. The repeat domains were identified visually and using the Internal Repeat Finder[2] algorithm. Protein secondary and tertiary structure was predicted using Protein Homology/analogY Recognition Engine V 2.0 (Phyre^2)[3]. Sequence alignment and protein identity was calculated with Clustal Omega algorithm program[4]. N-terminus signal sequence was predicted using SignalP 4.1 server[5]. Hydrophobic transmembrane domain was predicted using TMHMM Server v. 2.0[6].

Construction and purification of histidine-tagged fusion proteins

rSesJ$_{258\text{-}637}$ protein was expressed with hexa-histidine tag at the N-terminus using the expression vector pQE30 (Qiagen). PCR primers used were SdrS N2N3 fwd (TAG**GGATCC**CCAGAGGTTGATTCCGAAGTATTAG) and SdrSA rev

---

[2] http://nihserver.mbi.ucla.edu/Repeats/

[3] http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

[4] http://www.ebi.ac.uk/Tools/msa/clustalo/

[5] http://www.cbs.dtu.dk/services/SignalP/

[6] http://www.cbs.dtu.dk/services/TMHMM/

(TAG**GTCGAC**CTAAAGTTTTTCATTGCCAGTAGCAAC). Genomic DNA from *S. epidermidis* strain 3094 was isolated using UltraClean® Microbial DNA Isolation kit (MO BIO Laboratories, Inc.). Expression cultures were induced with IPTG and protein was purified using nickel-affinity chromatography and anion-exchange chromatography as described previously [28]. Binding buffer (20 mM sodium phosphate, 1 M NaCl, pH 8.0), wash buffer (20 mM sodium phosphate, 1 M NaCl, 25 mM Immidazole) and elution buffer (20 mM sodium phosphate, 1 M NaCl, 500 mM Immidazole, pH 8.0) were used for nickel-affinity chromatography. Binding buffer (20 mM Tris, pH 8.0), wash buffer (20 mM Tris, 50 mM NaCl, pH 8.0) and elution buffer (20 mM Tris, 500 mM NaCl, pH 8.0) were used for anion-exchange chromatography.

Flow cytometry

To determine surface expression of SesJ by flow cytometry, bacteria were grown for 4 h in TSB broth from an inoculum with an OD600 of 0.03. Collected and washed cells were labeled with preimmune or SesJ antiserum followed by Alexa Fluor 488 conjugated goat anti-rabbit IgG as described previously. Cells were fixed with 3% paraformaldehyde in PBS and analyzed with BD Accuri™ C6 cytometer. Polyclonal antibodies against $rSesJ_{258\text{-}637}$ were raised in rabbit (Rockland Immunochemicals Inc.). Pre-immune sera was tested for reactivity to SesJ using ELISA before selecting an animal for antibody production.

Sequence logo

Sequence logo was generated using frequency plot at WebLogo online program[7] [167, 168]. Custom color scheme as described in the figure legend was used to generate the graph.

PCR screening

Ninety five coagulase-negative *Staphylococcus* colonizing and infections isolates were used to determine the distribution of the *sesJ* gene, including 64 isolates from a study on the epidemiology of *S. epidermidis* colonization and infection in left ventricular assist device individuals [169]. DNA from picked colonies was amplified using primers SdrS-F 5′-GAGCACAGACAATTCGACTTCAAATC and SdrS-R TCAGCATATTCCGGCATATCTACTG and PCR products were sequenced for confirmation.

Results

Sequence analysis of the SesJ protein

While examining the sequence variation of SdrG in published *S. epidermidis* genomes, we discovered a gene encoding a SdrG-like but clearly distinct CWA protein that we have called SesJ (GenBank accession number: KU935462) according to the established nomenclature. The deduced full length SesJ protein in strain 3094 is 1047 amino acids

[7]http://weblogo.berkeley.edu/logo.cgi

long (Figure 11A). The predicted molecular mass of the mature SesJ protein after cleavage of the signal sequence and processing by sortase is 105.82 kDa. Amino acid sequence analysis of the full length protein revealed that SesJ has 41 and 40% sequence identity to SdrG (GenBank accession number: AAF72510.1) and SdrF (GenBank accession number: AAF72509.1), respectively. SesJ is a multidomain protein that contains, starting from the N-terminus, a 44 amino acid long signal sequence, a NTR region, an A-region, two B repeats, an aspartic acid containing repeat (ACR) region and typical cell wall anchoring sequences such as a LPXTG motif, a hydrophobic membrane spanning region and a short cytoplasmic positively charged tail (Figure 11A).

SesJ is a novel chimeric MSCRAMM

MSCRAMMs bind to their ligands through the DLL mechanism [87, 88]. This binding mechanism involves characteristic structural features in the MSCRAMM A-region including two adjacent IgG-like folded domains where a conserved TYTFTDYVD-like motif is present at the "back" of the latching trench in the first domain in the tandem and a latch sequence at the C-terminal extension of the second domain. A latch sequence is not a conserved sequence of amino acids but consists of small uncharged, polar and non-polar residues [87]. Tertiary structure prediction using the PHYRE^2 fold recognition server [170] indicated that residues 269–634 in the A-region of SesJ are highly likely to adopt two IgG-like folds (100% confidence level). The predicted structure of the two IgG-like domains in SesJ is very similar to the crystal structure determined for the N2N3

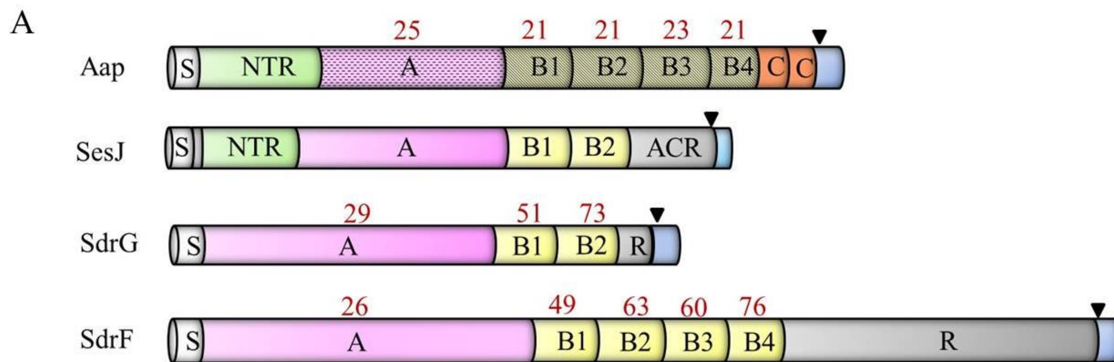


Figure 11. Comparison of SesJ with other structurally related proteins.

(A) Schematic representation of *S. epidermidis* surface proteins SesJ, SdrG, SdrF and Aap. Cartoons show the relative position of NTRs (SesJ and Aap) to other domains. The number in red above the individual domains represents its identity to the corresponding SesJ domain. The signal sequence is shown in white, NTRs in green, A-region in pink, B repeats in yellow, R region and ACR region in gray, collagen triple helix in orange, and cell wall spanning region and cytoplasmic tail in blue. LPXTG motif is shown using a black triangle. (B) Modeled 3D structure of the A-region of SesJ comprising two IgG-like folded domains compared to the crystal structure of $SdrG_{N2N3}$. (C) Structure predictions of B1 (yellow), B2 (green) repeat of SesJ, along with an overlay of SesJ B1 (yellow) and B2 (green), and SesJ B1 (yellow) with SdrD B1 (orange) repeat.

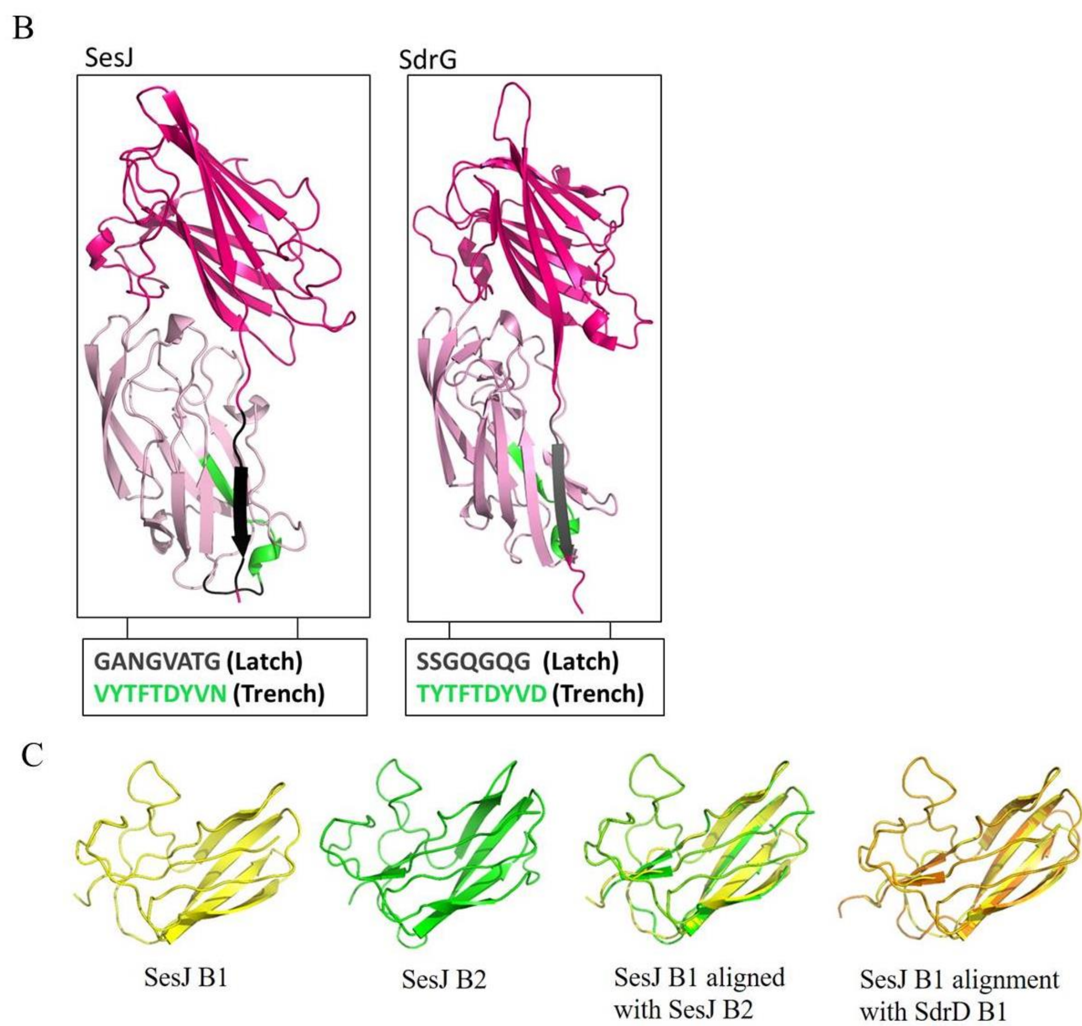


Figure 11. Continued.

Protein	Trench motif*	Latching sequence
SesJ	VYTFTDYVN	GANGVATG
ScsJ	VYTFTDYVN	GANGVATG
ShsJ	KYTFTDYVN	GANGIAQG
SdrI	TYTFTNYVD	GSSTAQG

* Residues in red indicate conserved residues amongst the proteins

Table 4. Conserved motifs in the A-region of SesJ structural homologs.

domain of SdrG [87] (Figure 11B). This SesJ segment furthermore contains the sequence, VYTFTDYVN at the expected positions for a latching trench in the first domain of the predicted IgG-like tandem. In addition, a putative latch sequence GANGVATG is present in the extension of the second IgG-like folded domain of SesJ. The presence of a latch sequence and a conserved TYTFTDYVD-like motif in the N2N3 sub-domain indicates that SesJ could bind a ligand peptide by the DLL mechanism (Table 4). In summary, the A-region of SesJ contains the characteristic IgG-like folded tandem of a MSCRAMM.

As a result of the unusually short N1 region that connects the N2N3 tandem to the preceding NTR region, the A-region of SesJ (378 amino acids) is smaller compared to the A-regions of SdrG (546 amino acid) and SdrF (624 amino acid). Sequence comparison shows that the amino acid sequence of the SesJ A-region is 29 and 26% identical to the comparable A-region segments of SdrG and SdrF, respectively (Figure11A). Similar levels of sequence identity are also seen between the A-regions of SesJ and the various *S. aureus* MSCRAMMs.

Staphylococcal MSCRAMMs do not usually contain a NTR domain, which is found in SesJ between the signal sequence and the A-region (Figure 11A). The NTR domain of SesJ is composed of a 15 amino acid long sequence repeated 10 – 13 times depending on the strain (Table 5). A NTR domain is also found in one other CWA protein from *S. epidermidis*; Aap, where it is composed of a variable number of a 16 amino acid long

Protein	Sequence of NTR*	Number/Length of repeat sequence
SesJ	-----EAPSKEEAPSNEATN	13/15
ScsJ	-----EEPSKEEATSKEVTN	8/15
ShsJ	-----EQASTEEKADTT---	24/12
SdrI	-----EPATKEEAATTE---	23/12
Aap	EAPQSEPTKTEEGSNA----	12/16

* Residues in red indicate conserved residues amongst the proteins

Table 5. Summary of NTRs of SesJ structural homologs.

sequence (Table 5). There is a low level of sequence identity throughout the Aap (GenBank accession number: AAW53239.1) and SesJ proteins (Figure 11A). The NTRs of the two proteins show intriguing similarities (further discussed below) but structure prediction show that the A-region of Aap is likely to adopt a lectin type fold, which is distinctly different from that of the characteristic MSCRAMM tandem [76].

SesJ contains two B repeats that are composed of 110 – 111 amino acids, similar in length to B repeats of the Sdr subfamily of staphylococcal MSCRAMMs. The SesJ B repeats show 63–69% sequence identity to B repeats of SdrG and SdrF (Figure 11A). Similar to SdrG and SdrF, the B repeats of SesJ also harbor predicted Ca^{2+} binding sites [171]. The predicted structures of SesJ B1 and B2 are shown (Figure 11C) and these are similar to the crystal structure determined for the *S. aureus* SdrD B1 [172]. On the other hand the crystal structure of the Aap B domain, which consists of two sub-domain G5 and E [173, 174] is very different from those predicted for the SesJ B domains.

The ACR region of SesJ is distinct from the SD repeated dipeptide characteristic of the Sdr protein subfamily of staphylococcal MSCRAMMs. A 20 amino acid long motif [SESTSESDSESHSDSES(H/D)SD] is repeated in the ACR region of SesJ. This type of arrangement is similar to the SRRs of SRRPs, which also are composed of longer motifs, e.g., SAS(T/E)SASTSASV in *S. gordonii* Challis [51]. We propose to name this segment ACRs since structural homologs of SesJ found in other CoNS species (see below)

contain similar C-terminal repeat motifs where an aspartic acid (rather than serine) is the conserved residue.

These comparative analyses reveal that SesJ is a new unique chimeric protein that has acquired structural motifs from other families of CWA proteins. The NTR are related to the corresponding segment of Aap, the IgG-like folded domains have all the features of a MSCRAMM and the ACR region resembles the SRR of streptococcal SRRPs.

SesJ is expressed on the surface of S. epidermidis

The deduced amino acid sequence of SesJ suggests that this protein could be expressed on the bacterial surface as a cell wall anchored protein. We verified by flow cytometry that SesJ in fact was expressed on the surface of *S. epidermidis* strain 3094 when bacteria were grown to mid exponential phase, 4 h (Figure 12A). *S. epidermidis* strain 2111, which does not have the *sesJ* gene, was used as a negative control to study expression of SesJ. The SesJ antisera did not cross react with any surface protein on the *S. epidermidis* 2111 strain. Similarly, pre-immune sera and secondary antibody alone did not bind to surface molecules on the *S. epidermidis* 3094 strain (Figure 12A). We next generated a growth curve (Figure 12B) and looked for expression of the protein at different time points during the growth curve. SesJ was constitutively expressed on the surface of *S. epidermidis* strain 3094 throughout the growth curve (Figure 12C).

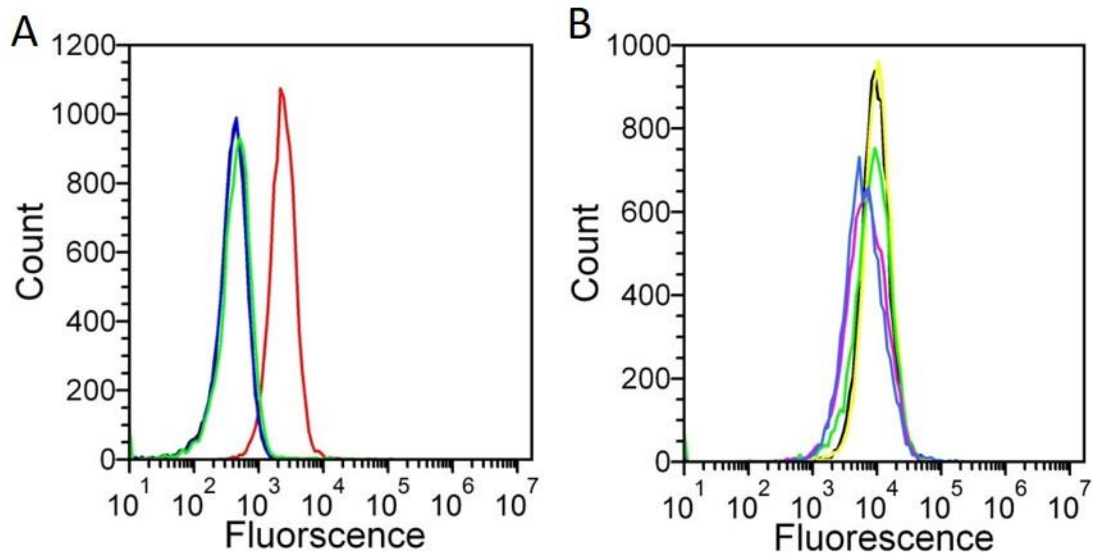


Figure 12. Detection of SesJ expression on the surface of *S. epidermidis* by flow cytometry. (A) Comparison of SesJ expression on the surface of *sesj*$^+$ *S. epidermidis* 3094 (solid red line), *sesj*$^-$ *S. epidermidis* 2111 (solid green line), *S. epidermidis* 3094 treated with pre bleed serum and secondary antibody (solid black line) and *S. epidermidis* treated with secondary antibody only (solid blue line)**.** (B) Growth curve of *sesj*$^+$ *S. epidermidis* strain 3094 in red line. Three hours is represented by black solid circle, 5 h by yellow solid circle, 8 h by green solid circle, 12 h by magenta solid circle and 25 h by blue solid circle. (C) Comparison of SesJ expression on the surface of *sesj*$^+$ *S. epidermidis* strain 3094 at 3 h (solid black line), 5 hrs (solid yellow line), 8 h (solid green line), 12 h (solid magenta line), 25 h (solid blue line).

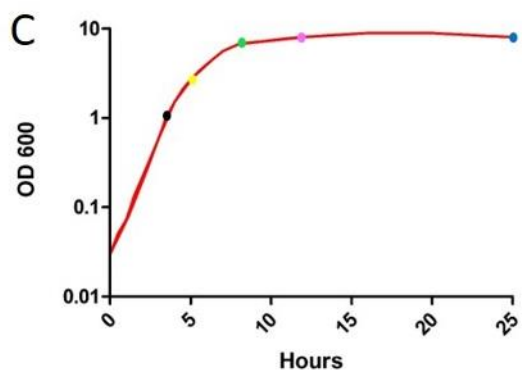


Figure 12. Continued.

Structural homologs to SesJ are present in other CoNS

We then searched for structural homologs in available sequenced genomes in the NCBI genome database for ORFs that contained NTRs and ACR region. Search results were further sorted for the presence of an A-region, B repeats and a LPXTG motif. We identified structural homologs of SesJ in *S. capitis*, *S. haemolyticus* and *S. saprophyticus* (Figure 13A). Two of these had not been previously identified and we have named these proteins ScsJ (*S. capitis* Strain CR01, GenBank accession number: WP_016898462.1) and ShsJ (*S. haemolyticus* strain JCSC1435, GenBank accession number: BAE03349.1). We found that the structural homolog in *S. saprophyticus* has already been reported and named SdrI (*S. saprophyticus* strain 7108, GenBank accession number: AAM90673.1) [175]. All these SesJ structural homologs contain a NTR region, an A-region, two B repeats and an ACR region as well as the characteristic cell wall anchoring motifs. The predicted lengths of the proteins vary considerably from 1048 to 1893 residues due to variations in the NTR and ACR regions.

The A-regions of these proteins are similar in length and range from 378 amino acid to 420 amino acid (Figure 13A). Tertiary structure prediction using the PHYRE^2 fold recognition server [170] indicated that residues in the A-region of the identified structural homologs is highly likely to also adopt the MSCRAMM characteristic IgG-domain contains a putative latch sequence (Table 4). The B repeats, which are of the Sdr protein types are similar in size (range from 108 amino acid – 120 amino acid) and contain putative Ca^{2+} binding sites.

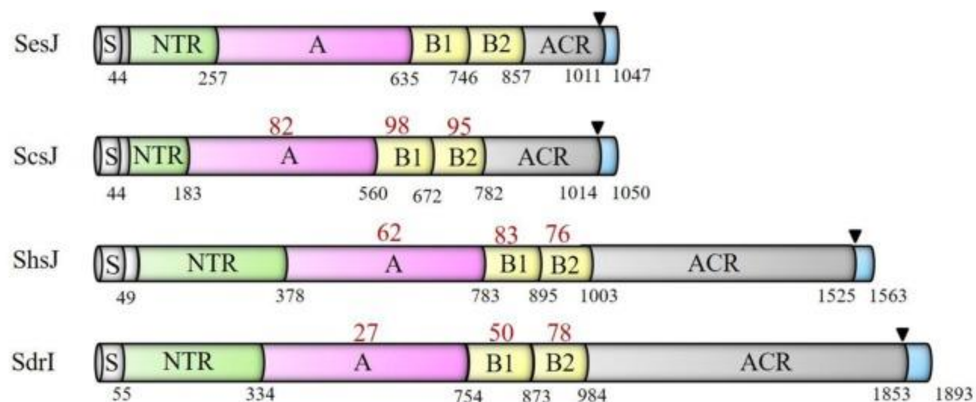


Figure 13. SesJ structural homologs in other CoNS species. (A) SesJ structural homologs have the same relative position of structural domains. The number in red below the individual domains represents its identity to the corresponding SesJ domain. Numbers in black represent the end of domain. Signal sequence is shown in white, NTRs in green, A-region in pink, B repeats in yellow, ACR region in gray, and cell wall spanning region and cytoplasmic tail in blue. LPXTG motif is shown using a black triangle. (B) Predicted structures of $ScsJ_{239\text{-}559}$, $SdrI_{440\text{-}755}$, $ShsJ_{462\text{-}782}$ and overlay of $SesJ_{269\text{-}634}$ (blue) with $ScsJ_{239\text{-}559}$ (green).

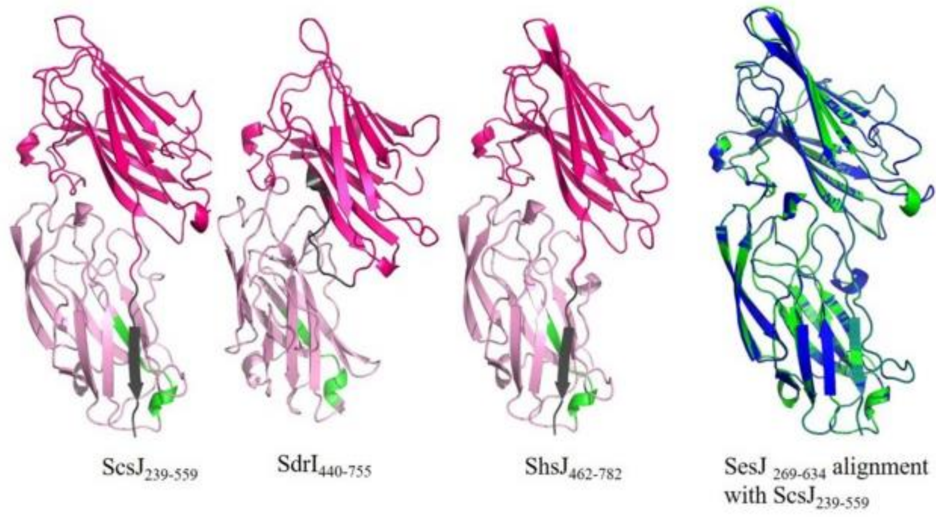


Figure 13. Continued.

like folded tandems (100% confidence; Figure 13B). The conserved TYTFTDYVD-like motif is found in the first IgG-like domain and the extension of the second IgG-like

SesJ is most similar to ScsJ from *S. captis* with sequence identity of 82 and 96% for the A-region and B repeats, respectively. The A-region of SesJ has about 62% identity with ShsJ A-region but less than 30% identity to that of SdrI. The B repeats are relatively well conserved in all identified members of this subfamily. B2 repeats are most conserved with over 75% identity whereas the B1 repeats show slightly more variations (up to 50%) (Figure 13A). This suggests that the A-region of these proteins may interact with different ligands.

The repeats in the ACR region of the proteins vary in amino acid composition and length of the repeat motif. ScsJ has the same residues in the ACR region as SesJ and a similar motif, i.e., SESESESHSDSESHSDSEST. ShsJ has a 12 amino acid repeat motif S(T/Q)SDSES(T/Q)SDSE which is shorter than repeat motifs in the ACR region of SesJ and ScsJ. SdrI has both serine-aspartic acid dipeptide and alanine-aspartic acid dipeptide repeats with a repeat motif $SD_{(1-2)}AD_{(1-5)}$.

The NTRs contain a conserved motif

The presence of NTRs defines this MSCRAMM subfamily but the number, length and sequence of the individual repeat units vary among the identified subfamily members (Table 5). The repeat units are rather long; 12–15 residues and are composed of mostly

hydrophilic amino acids. Within NTRs of individual proteins, the sequence in the first and last repeat units diverges somewhat from the sequence of repeats in the core of the NTRs. Despite the extensive variation among the sequences of different NTRs, an eight amino acid long conserved motif can be identified (Figure 14). Intriguingly the 16 amino acid long repeat unit in the NTR of Aap also contain a variant of the core motif [E***(K/T)EE] found in NTRs of SesJ homologs (Table 4).

Prevelence of SesJ structural homologs in CoNS

We examined 95 *S. epidermidis* isolates to determine the prevalence of the *sesJ* gene. Seventeen isolates (18%) were positive and the *sesJ* gene was present in both colonizing and infectious isolates. Furthermore, 4/26 (15%) *S. capitis* bloodstream isolates were positive for the *scsJ* gene. These frequencies are similar to the reported presence of *sdrI* (11%) in *S. saprophyticus* UTI samples [175]. Thus, members of the new subfamily of MSCRAMMs are present in 11–18% of the examined CoNS isolates.

Discussion

Pan genome sequence analysis has revealed the plasticity of the CoNS species genomes [18, 109]. These include CoNS genome changes by the addition of new genes, which can alter the virulence potential of the different subspecies. Sequencing of more genomes aids in the identification of new virulence factors. In this study, we identified and characterized a novel MSCRAMM SesJ, which is a chimeric protein with structural

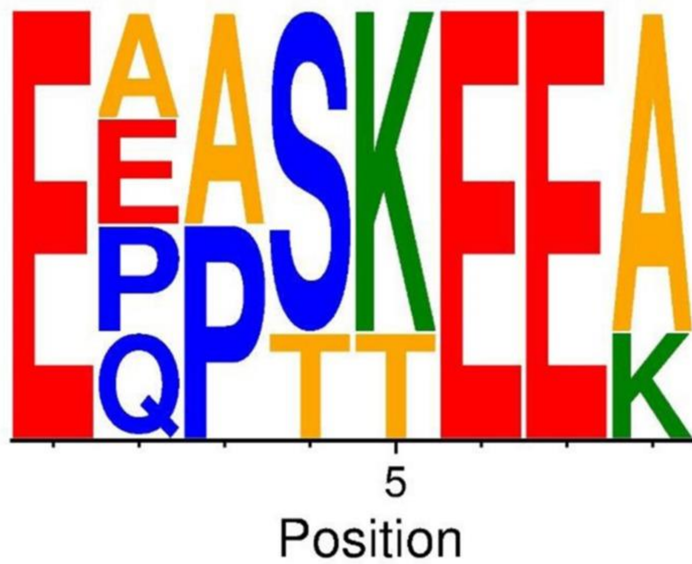


Figure 14. Conserved motif in the NTRs of SesJ structural homologs. Sequence logo of conserved 8 amino acid motif amongst all the structural homologs. Basic residues are colored in green, the acidic residues in red, the neutral ones in blue, the hydrophobic ones in yellow and remaining residues in orange.

features characteristic of staphylococcal MSCRAMMs. The NTR and ACR regions of SesJ are not found in other MSCRAMMs but are related to similar domains in Aap and SRRPs, respectively. Furthermore, SesJ is a prototype of a new subfamily of MSCRAMMs. This newly identified family consists of the structural homologs SesJ, ScsJ, ShsJ and SdrI, present in ~11–18% of *S. epidermidis*, *S. capitis*, *S. haemolyticus*, and *S. saprophyticus*, respectively. We also verified that SesJ is expressed constitutively on the surface of *S. epidermidis*.

Protein families based on structural homology have been reported before. Many gram positive bacteria express members of a family of structurally related collagen binding proteins where CNA of *S. aureus* is a prototype [102]. Similarly, the SRRPs constitute a family of structurally related proteins present in different streptococcal and staphylococcal species [51]. Structural homologs of Bap are present in *S. epidermidis*, *S. chromogenes*, *S. hyicus*, *S. xylosus*, *S. simulans* as well as in *S. aureus* [43].

The *S. aureus* MSCRAMMs have been studied in most detail. Although these proteins all have sequence features required for the structural organization of MSCRAMM their overall sequences can vary significantly [28, 77]. These sequence variations allow individual MSCRAMMs to have different functions, target different ligands or different sequences in the same ligand. For example, fibrinogen is a ligand for many of the *S. aureus* MSCRAMMs but ClfA, ClfB and Bbp bind to different sites in the fibrinogen molecule. The sequence variations seen among the SesJ like proteins in different CoNS

suggests that these proteins do not necessarily target the same ligand or even have the same function in the different species.

The C-terminal repeat regions of Sdr proteins and SRRPs are hypothesized to extend the ligand binding domain, i.e., the A-region of Sdr proteins and NR region of SRRPs, away from the cell surface to prevent obstruction of ligand binding by the cell wall [51, 78]. The C-terminal repeat regions of Sdr proteins and SRRPs are often glycosylated. SRRPs are encoded at loci harboring genes for two glycosyltransferases. Sdr proteins also have two glycosyl transferase *sdgA*, *sdgB* gene encoded downstream of SdrC/D/E proteins [176]. Likewise, two genes predicted to encode glycosyl transferases are located immediately upstream of the *sesJ* gene (not shown) and we therefore speculate that SesJ is glycosylated. These observations indicate that SesJ is subjected to similar post translational modifications as the Sdr proteins and the SRRPs.

Cell wall anchored proteins of gram positive bacteria contain a characteristic C-terminal region where the LPXTG motif is recognized by the transpeptidase sortase, which covalently attaches the protein to the peptidoglycan on the surface of the organism. Flow cytometry confirmed that SesJ is present on the surface of *S. epidermidis*. Furthermore, the protein is present from early logarithmic to late stationary phase in strain 3094. In *S. aureus*, expression levels of MSCRAMMs can vary during the growth phase. For example, expression of FnbpA on the surface of *S. aureus* Newman increases during late logarithmic phase and decreases in the late stationary phase. In contrast, ClfB increases

during early logarithmic phase and disappears at early stationary phase [177]. The expression of SdrG on the surface of *S. epidermidis* strain 0-47 was not induced under the tested in vitro growth conditions but increased over 30 fold in 3 h in a murine infection model [178]. Thus the expression of Sdr proteins on staphylococci can differ, depending upon the nature of the protein and the environment, and these variations may relate to the role of the proteins in the pathogenic mechanism of the organism. Additional studies of the expression of SesJ in different strains and under different conditions are warranted and can provide clues on the role of the encoded protein in the life of *S. epidermidis*.

The A-region is the primary ligand binding region of MSCRAMMs, which often uses the DLL mechanism to engage the targets. $SdrG_{N2N3}$ binds to a specific sequence in the N-terminal section in the β-chain of human fibrinogen by this mechanism [29, 88]. However, B repeats may also engage in ligand binding. For example SdrF binds Type I collagen via its B repeats in a temperature dependent manner [179]. MSCRAMMs bind to host proteins and can mediate bacterial adhesion, evade immune response as well as participate in biofilm formation. Both SdrG and SdrF initiate biofilm formation by binding to human fibrinogen and Type I collagen, respectively [53, 87]. Aap participates in both the initial attachment phase, where a region containing the NTRs and the A domain is required, and in the accumulation phase, where the B domains seem to be involved [174, 180]. SRRPs acts as adhesins and colonize host tissues by forming biofilms. SesJ's similarity to MSCRAMMs and domains from proteins involved in

biofilm formation, leads to a speculation that SesJ might be involved in biofilm formation and/or interact with host proteins. The interactions of SesJ and its possible role in *S. epidermidis* pathogenesis are currently under investigation.

CHAPTER III

SESJ AND PLS ARE PRESENT IN A COMPOSITE SCC*mec* ELEMENT

Introduction

S. epidermidis is an opportunistic pathogen associated with infections in patients with an implanted foreign body or a compromised immune response [2, 7, 15]. With advances in medicine, the number of patients with both an implanted foreign body and a suppressed immune system has increased. These patients carry a high risk for infections and the increased misuse of antibiotics has exacerbated this condition. *S. epidermidis* has successfully thrived in this ecological niche, establishing itself as a successful nosocomial pathogen. *S. epidermidis* undergoes frequent genetic recombination leading to well-adapted clonal lineages e.g. CC2 [181, 182]. Using MLST, the genetic background of infectious isolates has been identified around the globe. These studies revealed that multiple sequence types (STs) are capable of causing infection. ST2, the founder sequence type of clonal complex 2, is the most predominant and geographically widespread clone [181-183]. A high frequency of genetic recombination within CC2 has resulted in an unusually large number of sequence types in this genetic lineage. These studies also provided clues to the epidemic nature of *S. epidermidis* [50, 182, 183]. Perhaps, *S. epidermidis* has been successful as an opportunistic pathogen because of its capability of high frequency genetic recombination and gene acquisition, especially the acquisition of antibiotic resistance determinants.

70-90% of *S. epidermidis* clinical isolates are methicillin resistant and carry the *mecA* gene present in the SCC*mec* element [50, 183]. The SCC*mec* elements carrying the *mecA* gene are widespread among coagulase negative staphylococci [184-187]. To date, XI SCC*mec* elements have been identified in staphylococci, 8 of which are found in the human isolates [138]. SCC*mec* elements have four characteristic features (i) the presence of methicillin resistance gene *mecA*, (ii) the presence of cassette recombinases ccrA/B or ccrC (iii) insertion at integration site sequence *att* in the *orfX* gene and (iv) the presence of flanking direct repeats on both sides of the integrated SCC*mec* [185]. SCC*mec* Type IV is the most frequent SCC*mec* represented in *S. epidermidis* clinical isolates [180, 181]. In addition, SCC*mec* Type I, II, III, V, VI, VII also have been identified in *S. epidermidis* clinical isolates collected from different patient types and environments [189-191].

Genetic recombination has led to multiple novel configurations of the SCC elements deviating from the identified types and sub-types. Pseudo elements missing either the recombinase genes (ψSCC*mec*) or a functional *mec* gene complex have been described [188, 192-194]. Additionally, some staphylococci isolates contain more than one SCC*mec* element [184, 186]. A study by Zong et. al. reported that 23 out of 84 methicillin resistant CoNS had more than one SCC*mec* element [184]. Similarly, SCC*mec* elements with two set of recombinase genes have been identified [195]. Lastly, some SCC non-*mec* elements contain virulence genes instead of the *mecA* gene. For example, SCC*cap1* contains the genes for capsular polysaccharide while SCC*Hg* carries

the mercury resistance operon [196-198].

CWA proteins are critical for establishing a bacterium as a commensal and a pathogen [28]. CWA proteins are in direct contact with the host. Hence, it is not surprising that CWA proteins have multiple functions e.g. interact with host proteins, promote bacterial survival and evade the host immune response to establish infection. The MSCRAMM family is one of the well-characterized CWA protein families in staphylococci. MSCRAMMs contain two tandemly linked IgG-like folded domains required for binding the ligand by DLL mechanism [87, 88, 107, 199, 200]. Many MSCRAMMs are critical virulence factors involved in adhesion, invasion, and immune evasion of the bacteria. These functions are critical for establishing an infection in the host [30, 31, 33, 80, 89, 103]. We recently reported on the new NTR containing subfamily of MSCRAMMs present in coagulase negative staphylococci. *S. epidermidis* surface protein SesJ is the prototype of the new NTR containing subfamily of the MSCRAMMS [201].

Bacteria can adapt to selective pressure by acquiring MGE encoding antibiotic resistance determinants and virulence factors [43, 110]. Cell wall anchored proteins present in MGE provide selective advantage to bacteria [119]. For example, Plasmin sensitive protein (Pls) is present in SCC*mec* Type 1 [202, 203]. Pls plays a role in biofilm formation through the G5 repeats as well as the glycosylated serine aspartate dipeptide repeat region [204, 205]. In a mouse septic arthritis model, the presence of the *pls* gene

also led to more frequent joint infection and severe arthritis [206]. SesJ protein is only present in the 18% of the previously tested *S. epidermidis* clinical isolates [201]. We hypothesized that the *sesJ* gene is present in a mobile genetic element. Here we report the presence of *S. epidermidis* cell wall anchored protein SesJ in a composite SCC*mec* element. We also discovered that the *pls* gene is present in SCC*mec* Type IV in *S. epidermidis* clinical isolates.

Materials and methods

Isolates

Serial isolates from blood were collected for the cancer patients that presented with symptoms of bloodstream infections at MD Anderson hospital, Houston, Texas. *S. epidermidis* isolates were confirmed using MALDI-TOF.

Multilocus sequence typing and clonal complex determination

The draft genome assemblies of the isolates were utilized to determine the MLST by batch sequence query of the Bacterial Isolate Genome Sequence Database (BIGSdb) [207]**.** Clonal complexes were determined by the eBURST algorithm, and displayed by PHYLOViZ software v2 [208]. The most exclusive group definition used for eBURST algorithm included STs within the same group if they share identical alleles at six or seven of the seven MLST loci with at least one other ST in the group.

Library preparation

Genomic DNA for Illumina sequencing was extracted using the MasterPure Kit (Illumina Inc., San Diego CA) from cultures grown overnight in Muehller-Hinton (MH) broth at 37º C. 10 µg of gDNA was used for paired-end sequencing on the Illumina® instrument (Illumina, Inc. San Diego, CA) using TruSeq chemistry at the MD Anderson Sequencing and Microarray Facility. Genomic DNA for Pacbio sequencing was extracted using phenol/chloroform extraction from overnight cultures. The Pacific Biosciences guidelines for preparing 20 kb SMRTbell template were used to create a large insert library.

Genome Sequencing

Short read quality for reads generated through Illumina sequencing was assessed using the FASTQC toolkit (Babraham Bioinformatics). Adaptors as well as low quality reads were trimmed using Trimmomatic v 0.33 [209]. Genome assembly for short reads was performed using SPAdes v 3.9.1 [210]. 28 genomes assemblies either failed, or were deemed poor (> 300 contigs) using SPAdes assembly, were assembled with SeqMan Ngen v 14.1 (DNASTAR® Inc, Madison WI). Custom scripts were used to determine the average depth of SPAdes generated draft genomes, and remove contigs with less than 300 bp and the phi X 174 sequence. This pipeline was implemented on the high performance computing cluster at MD Anderson. HGAP pipeline was implemented for de novo genome assembly of PacBio reads [211]. PacBio assembly results were corrected by mapping paired-end Illumina short reads using Bowtie2 v 2.2.3 [212].

SCCmec typing

SCC*mec* type was determined *in silico* by typing the combination of *mec* complex class and the type of *ccr* complex. Different *mec* gene complexes and *ccr* gene combinations were queried in the assembled genome sequence data of all the study isolates. The presence of the *mec* gene complex and the *ccr* gene complex was further confirmed by visualization of read alignment across the entire genome. A SCC*mec* element was considered typeable if it contained a previously known combination of the *mec* gene complex and the *ccr* gene complex.

Determination of presence of virulence genes

Known *S. epidermidis* virulence factor gene or protein sequences were queried in the assembled genome sequence data of the study isolates. Protein sequence for SesJ (accession number ANG65522.2), Pls (accession number BAA86640.2), Ebh (accession number AAW55290.1), Aae (reference sequence number NP_765874.1), Aap (accession number AAW53239.1), SdrF (accession number AAF72509.1), SesC (accession number AAW53125.1), AltE (accession number AAW53968.1), SdrG (accession number AAF72510.1), SesI (accession number AAW54982.1) were used to query the assembled genomes. Nucleotide sequence for *icaRADBC* operon (accession number U43366.1), Type 1 *agr* operon (accession number Z49220.1), Type 2 *agr* operon (accession number FJ707317.1), Type 3 *agr* operon (accession number CP009046.1; region 607512..611854) were used to query the assembled genomes. The presence of the gene was further confirmed by visualization of read alignment to the virulence gene.

MB ID	ST	mecA	Agr	ica RADBC	SCCmec	sesJ	pls	ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB1048	2	Y	1	Y	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1119	2	Y	1	Y	NT1	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1145	2	Y	1	Y	NT3	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1153	2	Y	1	Y	NT4	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1208	2	Y	1	Y	NT1	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1306	2	Y	1	Y	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1502	2	Y	1	Y	VII	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1567	2	Y	1	Y	NT4	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1569	2	Y	1	Y	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1595	2	Y	1	Y	IV	N	Y	Y	Y	ND	Y	Y	Y	Y	N
MB1617	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB1651	2	Y	1	Y	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1656	2	Y	1	Y	III	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1708	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB1721	2	Y	1	Y	VII	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB1750	2	Y	1	Y	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1985	2	Y	1	Y	ND	Y	N	Y	Y	ND	Y	Y	Y	Y	N
MB2095	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB247	2	Y	1	Y	NT4	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB323	2	Y	1	Y	ND	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB422	2	Y	1	Y	NT4	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB434	2	Y	1	Y	NT4	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB463	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB466	2	N	1	Y	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB469	2	N	1	Y	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB503	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB546	2	Y	1	Y	NT3	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB551	2	Y	1	Y	III	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB591	2	Y	1	Y	IV	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB609	2	Y	1	Y	ND	N	N	Y	Y	Y	Y	Y	Y	Y	Y
MB619A	2	Y	1	Y	III	N	N	Y	Y	ND	Y	Y	Y	Y	N

Table 6. Presence of virulence genes in the isolates.

MB ID	ST	mecA	Agr	ica RADBC	SCCmec	sesJ	pls	ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB918	2	Y	1	Y	VII	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB931	2	Y	1	Y	III	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB941	2	Y	1	Y	VII	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1335	6	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1464	6	Y	1	N	IV	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB496	6	N	1	Y	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB939	6	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB101	7	N	1	Y	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB539	7	N	1	Y	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1182	16	N	1	Y	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1230	16	N	1	Y	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1849	16	Y	1	Y	NT1	N	N	Y	Y	N	N	Y	Y	Y	N
MB601D	16	Y	1	Y	IV	N	N	Y	Y	Y	N	Y	Y	Y	N
MB888A	16	Y	1	Y	ND	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1229	20	Y	1	Y	NT1	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2050A	20	N	ND	ND	-	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB995	20	N	1	Y	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1091	22	Y	1	Y	III	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1619	22	Y	1	Y	III	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB724	22	N	3	Y	-	N	N	Y	Y	N	Y	Y	Y	Y	N
MB927	22	Y	1	Y	III	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB947	22	Y	1	Y	III	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB123	23	Y	1	Y	ND	N	N	Y	Y	N	N	Y	Y	Y	N
MB1915	23	N	1	Y	-	N	Y	Y	Y	N	N	Y	Y	Y	N
MB775	35	Y	1	Y	NT1	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1299	57	N	2	N	-	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1421	59	N	3	N	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1910	59	Y	1	N	IV	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB283	59	Y	1	N	ND	N	Y	Y	Y	Y	Y	Y	Y	Y	N
MB585B	59	Y	1	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB536	69	Y	2	Y	IV	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB714	69	Y	2	Y	NT5	N	N	Y	Y	Y	Y	Y	Y	Y	N

Table 6. Continued

MB ID	ST	mecA	Agr	ica RADBC	SCCmec	sesJ	pls	Ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB906	69	Y	2	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB113	72	N	1	Y	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB192	73	N	3	N	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB037	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1082	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1298	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1546	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1547	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1568	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1675	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1825	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1956	83	Y	ND		ND	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB268	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB449	83	N	1	Y	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB539A	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB56	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB599	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB821	83	Y	1	Y	IV	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB899	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB999	83	Y	1	Y	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2013	87	Y	2	N	IV	N	N	Y	Y	N	Y	Y	Y	Y	N
MB1043	88	Y	3	N	IV	N	N	Y	ND	Y	Y	Y	Y	Y	N
MB1260	89	N	1	N	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1609	89	N	1	N	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1364	130	Y	1	N	ND	N	N	Y	Y	N	Y	Y	Y	Y	N
MB316	130	Y	1	Y	IV	N	N	Y	Y	N	Y	Y	Y	Y	N
MB445	130	Y	1	N	IV	N	N	Y	Y	N	Y	Y	Y	Y	N
MB415	173	N	2	Y	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1457	190	Y	1	Y	ND	N	N	Y	Y	ND	Y	Y	Y	ND	N
MB1812	203	N	1	N	-	N	N	Y	Y	N	Y	Y	Y	Y	N
MB1143	210	Y	1	N	ND	Y	N	Y	Y	N	Y	Y	Y	Y	N
MB1586	210	Y	1	N	ND	Y	N	Y	Y	N	Y	Y	Y	Y	N

Table 6. Continued

MB ID	ST	mecA	agr	ica RADBC	SCCmec	sesJ	pls	Ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB1709	210	N	1	N	-	Y	N	Y	Y	N	Y	Y	Y	Y	N
MB1715	210	N	1	N	-	Y	N	Y	Y	N	Y	Y	Y	Y	N
MB1907	210	Y	1	N	VII	Y	N	Y	Y	N	Y	Y	Y	Y	N
MB526	210	Y	1	N	VII	Y	N	ND	Y	N	Y	Y	Y	Y	N
MB567	218	N	3	N	-	Y	N	Y	Y	Y	N	Y	Y	Y	N
MB312	228	N	3	N	-	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1830	329	N	3	N	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB429	520	N	1	N	-	N	N	Y	Y	Y	N	Y	Y	Y	N
MB1872	540	N	1	N	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB1772	5	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB896	5	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1057	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1096B	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1191	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1193	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1226	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1278	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1427	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB151C	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1593	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1669	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1769	5-A	Y	2	N	IV	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB1851	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1871A	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1999	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2033	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2126	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2177A	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB746	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB806	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB929	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N

Table 6. Continued

MB ID	ST	mecA	agr	ica RADBC	SCCmec	sesJ	pls	Ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB951	5-A	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB10	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1034	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1037	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1064	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB108	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1211	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1218	5-B	Y	2	N	IV	N	N	Y	Y	ND	Y	Y	Y	ND	N
MB1312	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1347	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1368	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1489	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB149	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1532	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1542	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB1679	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
MB1691 A	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
MB1695	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
MB1961	5-B	Y	2	N	ND	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB1962	5-B	Y	2	N	IV	N	N	Y	Y	ND	Y	Y	Y	ND	N
MB198B	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB2193	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB227A	5-B	Y	2	N	NT2	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB383	5-B	Y	1	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB396	5-B	Y	2	N	NT2	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB492	5-B	Y	2	N	NT2	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB524	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB606	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB634	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB716	5-B	Y	2	N	NT2	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB740	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB860	5-B	Y	2	N	IV	N	N	Y	Y	Y	Y	Y	Y	Y	N

Table 6. Continued

MB ID	ST	mecA	agr	ica RADBC	SCCmec	sesJ	pls	Ebh	Aae	Aap	sdrF	sesC	altE	sdrG	SesI
MB883	5-B	Y	2	N	NT2	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB905	5-B	Y	2	N	NT2	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB972	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB988	5-B	Y	2	N	IV	Y	N	Y	Y	Y	Y	Y	Y	Y	N
MB182	ND	Y	1	N	IV	N	Y	N	Y	N	N	Y	Y	Y	N
MB1908	ND	N	3	N	-	N	N	N	Y	N	N	Y	Y	Y	N
MB1935	ND	N	3	N	-	N	N	ND	Y	N	N	Y	Y	Y	N
MB255	ND	N	3	N	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB339	ND	N	3	N	-	N	N	Y	Y	N	N	Y	Y	Y	N
MB446	ND	Y	2	N	NT1	N	N	Y	Y	Y	Y	Y	Y	Y	N
MB497	ND	N	2	Y	-	N	N	ND	Y	N	Y	Y	Y	Y	N
MB89	ND	Y	ND	Y	NT1	N	N	Y	Y	ND	Y	Y	Y	Y	N
MB895	ND	N	2	N	-	N	N	Y	Y	N	N	Y	Y	Y	N

NT1 represents a pseudo SCC element containing the *mec* gene complex but lacking the *ccr* genes
NT2 represents a SCC*mec* Type IV element with an additional *ccrC8* enzyme gene.
NT3 represents a SCC*mec* Type VII element with an additional *ccrA4B4* gene complex.
NT4 represents a SCC*mec* Type VII element with an additional *ccrA4B5* gene complex.
NT5 represents a SCC*mec* Type IV element with an additional *ccrC2* enzyme gene.
ND stands for could not be determined.

Table 6. Continued

Results

Epidemiology study

We retrospectively collected 171 clinical isolates obtained from blood samples of cancer patients at MD Anderson hospital, Houston, Texas. 30 different ST and 9 isolates with unknown ST were identified in the collection (Table 6). ST 2, ST 5 and ST83 accounted for 64% of the total isolates. 58 of the isolates corresponded to ST5 (34%), 34 of the isolates corresponded to ST 2 (20%) and 17 isolates were identified as ST83 (10%). 6 of the isolates belonged to ST210. 16 of the STs were singletons, while ST6, ST7, ST16, ST20, ST22, ST23, ST59, ST69, ST89 and ST130 were each represented by 2-5 isolates. STs were grouped into two clonal complexes CC2 and CC5. In this study, CC2 comprised of founder ST2, ST22, ST 35 and ST16. CC5 comprised of founder ST5, ST6, ST130, ST210, ST7, ST190, ST87, ST59, ST1005, ST89 and ST88.

The *sesJ* gene is present in 30/171 (18%) of these isolates. This data is consistent with the previously reported epidemiology study performed using the blood isolates from patients with infected left ventricular assist drivelines [201]. The *sesJ* gene was detected in 6/34 of ST2 isolates (18%), 16/58 of ST5 isolates (26%) and 6/6 of ST210 isolates (100%) (Table 7). 2/16 singelton ST were also positive for *sesJ* gene. Additionally, we identified the *pls* gene in 10/ 171 (6%) isolates. 7/34 (21%) of ST2, 1/4 (25%) of ST59 and 1/2 of ST23 were identified as *pls*$^{+}$ isolates (Table 7). The *sesJ* and *pls* gene did not occur together in any of the isolates (Table 6). Both the *sesJ* gene and the *pls* gene are

ST	Number of isolates	*sesJ*$^+$ isolates (percentage)	*pls*$^+$ isolates (percentage)
ST 2	34	6 (18)	7(21%)
ST 83	17	0	0
ST 210	6	6 (100)	0
ST 5	58	16 (28)	0
ST23	2	0	1
ST59	4	0	1
Others	25	0	0
Singeltons	16	2	0
Unknown	9	0	1
Total	171	30 (17.5)	10 (6%)

Table 7. Presence of the *sesJ* and the *pls* gene in different ST.

present in selective ST which points towards genetic differences in the STs.

Distribution of MRSE and MSSE isolates

Isolates analyzed in this study included 139 (81%) MRSE isolates and 32 (19%) MSSE isolates. 32 MSSE isolates were distributed in over 20 different STs. However, ST2, ST5, ST83 i.e. STs with the most number of isolates, together comprised only 4 MSSE isolates (Table 6). The *sesJ* gene was present in 26/139 MRSE (19%) isolates and 4/32 MSSE (13%) isolates (Table 8). Out of 4 *sesJ*$^{+}$ MSSE isolates, 2 isolates belong to ST210, and other 2 isolates were singletons ST218 and ST57. Thus, the *sesJ* gene is present in both MRSE and MSSE isolates, with no preference for either one. On the other hand, the *pls* gene is present in the MRSE isolates only (Table 6 and 8).

Presence of virulence genes

The presence of *S. epidermidis* virulence related genes was tested in the isolate collection, which revealed ST related patterns (Table 6). For example, ST2 and ST83 isolates in this study encode for *ica* operon genes while ST210 and ST5 isolates do not. The *agr* quorum sensing system in *S. epidermidis* controls biofilm formation by regulating secreted protease production. ST2, ST83, ST210 isolates in the collection contain *agr* Type 1 genes whereas ST5 isolates (except for one) contain *agr* Type II genes. The *agr* type III genes were least represented in the collection. Ebh, SdrG, Aae, AltE, SesC, SdrF, Aap have all been shown to play a role in biofilm formation. At least 97% of the isolates contained the *ebh*, *sdrG*, *aae*, *altE* and *sesC* genes. 147/171 (86%)

	Number of isolates	***sesJ*** $^{+}$ **isolates (percentage)**	***pls*** $^{+}$ **isolates (percentage)**
MRSE	139	26(19)	10 (7)
MSSE	32	4(13)	0(0)

Table 8. Presence of the *sesJ* and *pls* gene in MRSE and MSSE isolates.

isolates contained the *sdrF* gene. 128/171 (75%) isolates were confirmed positive for the *aap* gene. Only 4 isolates encoded the *sesI* gene. Interestingly, all ST210 isolates contain the *sesJ* gene but not *icaADBC* and *aap* genes. In conclusion, presence or absence of certain genes correlated with the ST.

Distribution of SCCmec amongst the isolates

SCC*mec* types were assigned to 105 of 139 isolates while 33 isolates had untypeable SCC*mec* elements (Table 6). One isolate contained both the SCC*mec* element Type IV and Type VII. Out of XI SCC*mec* types, the isolate collection contained SCC*mec* Type III, Type IV and Type VII. 91 of 139 (65%) of MRSE isolates harbored SCC*mec* Type IV, followed by 8 (6%) isolates contained Type III and 6 (4%) isolated were assigned SCC*mec* Type VII. Amongst the untypeable isolates, 14 isolates had a combination of typeable SCC*mec* and additional *ccr* enzyme complex. Furthermore, 7 of the untypeable isolates had pseudo-SCC*mec* elements which lacked *ccr* genes. No other SCC*mec* types were detected.

SCC*mec* Type IV was present in a wide range of STs: ST2, ST83, ST59, ST69, ST130, ST6, ST16, and ST5. SCC*mec* Type III was harbored in ST2 and ST22 isolates. Additionally, SCC*mec* Type VII was represented in ST2 and ST210. ST2 had the most diversity in SCC*mec* with Type IV, Type VII, Type III, combination of two SCC*mec*, pseudo elements and SCC*mec* elements with the additional *ccr* enzyme complex. All SCC*mec* Type IV ST2 isolates either contained the *sesJ* gene or the *pls* gene. Whereas,

ST5 isolates harbored either Type IV or SCC*mec* elements with Type IV and additional *ccr* enzyme complex, and ST83 had only Type IV. As reported before, ST2 has a high recombination frequency, which is also evident here with the diversity of SCC*mec* elements.

SesJ and Pls are present in SCCmec elements

The *sesJ* and *pls* genes are both present in mobile genetic elements called SCC*mec* elements. In this study, MRSE blood isolates obtained from cancer patients encode the *sesJ* gene in SCC*mec* Type IV, Type VII and novel SCC*mec* elements containing additional *ccr* genes. Previous reports have shown that the *pls* gene is present in SCC*mec* Type I element [199, 200]. Whereas, in this collection the *pls* gene is only present in SCC*mec* Type IV element (Table 6).

The NCBI database search yielded in 4 *sesJ*$^{+}$ *S. epidermidis* isolates. *S. epidermidis* NIH06004 and *S. epidermidis* NIH04003 isolates were collected from blood of cancer patients with an infected catheter. *S. epidermidis* NIH051475 was collected from an infected joint of a patient with myelodysplastic syndrome. Lastly, *S. epidermidis* NIHLM039 was obtained from the retro auricular crease of a healthy individual. *S. epidermidis* NIH06004, NIH04003 and NIH051475 are methicillin resistant and encode the *sesJ* gene in the composite SCC*mec* elements containing remnants of SCC*mec* Type IV (Figure 15). Although *S. epidermidis* NIHLM039 lacks the *mecA* gene, it has other characteristic features of the SCC element i.e. the *ccr* gene, direct repeats and insertion

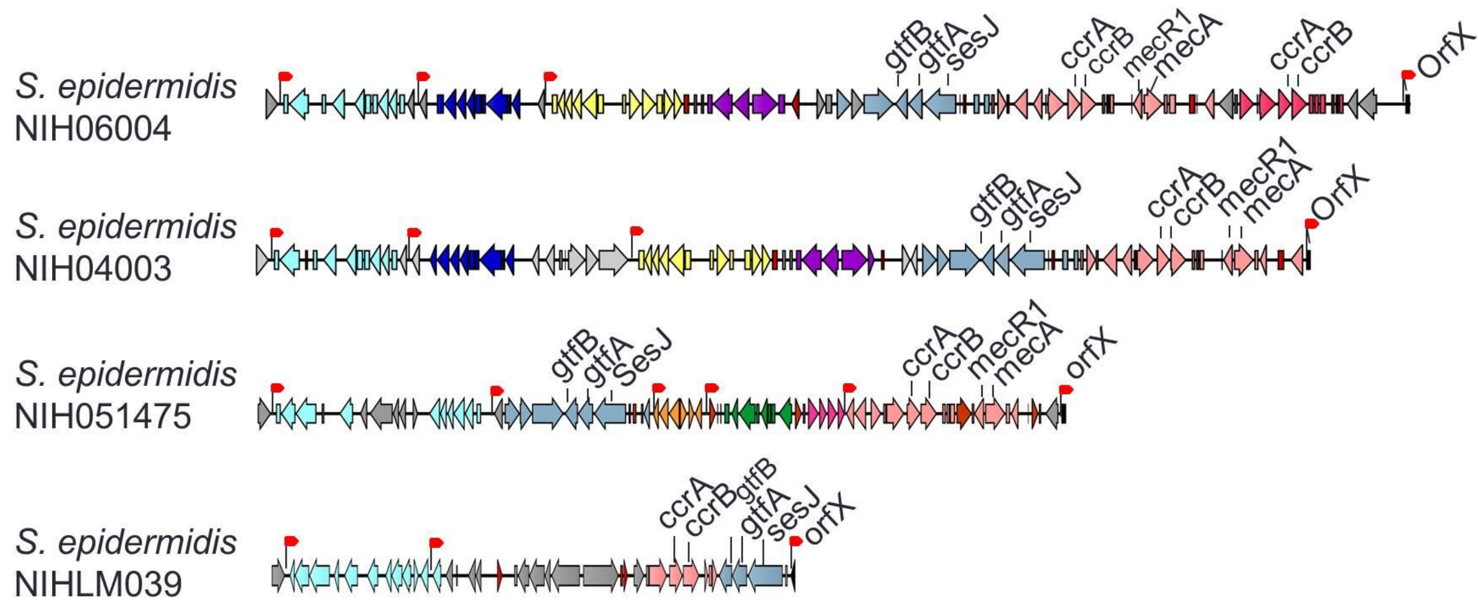


Figure 15. The *sesJ* gene is located in composite SCC*mec* element. Red flags represent presence of direct repeats.

```
ST2        PEVDSEVLDNSKQGTDNDNSSNHSEDNIVTYAEPTALTNTRSVDSPSRVSSDNSQAQKQG
ST57       PEVDSEVLDNSKQGTDNDNSSNHSEDNIVTYAEPTALTNTRSVDSPSRVSSDNSQAQKQG
ST218      PEVDSEVLDNSKQGTDNDNSSNHSEDNIVTYAEPTALTNTRSVDSPSRVSSDNSQAQKQG
ST5        PEVDSEVLDNSKQGTDNDNSSNHSEDNIVTYAEPTALTNTRSVDSPSRVSSDNSQAQKQG
ST210      PEVDSEVLDNSKQGTDNDNSSNHSEDNIVTYAEPTALTNTRSVDSPSRVSSDNSQAQKQG
           ************************************************************

ST2        KNVNDSIKVNSVDTDKEYVEPNNGQGFSTNVSFEVDGKVNKGDYFTVDMPEYADFNGIAD
ST57       KNVNDSIKVNSVDTDKEYVEPNNGQGFSTNVSFEVDGKVNKGDYFTVDMPEYADFNGIAD
ST218      KNVNDSIKVNSVDTDKEYVEPNNGQGFSTNVSFEVDGKVNKGDYFTVDMPEYADFNGIAD
ST5        KNVNDSIKVNSVDTDKEYVEPNNGQGFSTNVSFEVDGKVNKGDYFTVDMPEYADFNGIAD
ST210      KNVNDSIKVNSVDTDKEYVEPNNGQGFSTNVSFEVDGKVNKGDYFTVDMPEYADFNGIAD
           ************************************************************

ST2        YKAANNKIYPTINDGEQVVANGVYDTETKKLVYTFTDYVNKKDNIKGQFEIPQFIDRKNA
ST57       YKAANNKIYPTINDGEQVVANGVYDTETKKLVYTFTDYVNKKDNIKGQFEIPQFIDRKNA
ST218      YKAANNKIYPTINDGEQVVANGVYDTETKKLVYTFTDYVNKKDNIKGQFEIPQFIDRKNA
ST5        YKAANNKIYPTINDGEQVVANGVYDTETKKLVYTFTDYVNKKDNIKGQFEIPQFIDRKNA
ST210      YKAANNKIYPTINDGEQVVANGVYDTETKKLVYTFTDYVNKKDNIKGQFEIPQFIDRKNA
           ************************************************************

ST2        KTSGDYDLNYNIADKTVSKPMKIVYNNYDEGHVVANTSSLITKADLFNVGSHDYTQYIYV
ST57       KTSGDYDLNYNIADKTVSKPMKIVYNNYDEGHVVANTSSLITKADLFNVGSHDYTQYIYV
ST218      KTSGDYDLNYNIADKTVSKPMKIVYNNYDEGHVVANTSSLITKADLFNVGSHDYTQYIYV
ST5        KTSGDYDLNYNIADKTVSKPMKIVYNNYDEGHVVANTSSLITKADLFNVGSHDYTQYIYV
ST210      KTSGDYDLNYNIADKTVSKPMKIVYNNYDEGHVVANTSSLITKADLFNVGSHDYTQYIYV
           ************************************************************

ST2        NPKSEDSYNTRLTIQGYQEDVNDSSTLLNPDDTKIEILDAKSSDNIVPSFHINDEDFEDV
ST57       NPKSEDSYNTRLTIQGYQEDVNDSSTLLNPDDTKIEILDAKSSDNIVPSFHINDEDFEDV
ST218      NPKSEDSYNTRLTIQGYQEDVNDSSTLLNPDDTKIEILDAKSSDNIVPSFHINDEDFEDV
ST5        NPKSEDSYNTRLTIQGYQEDLNNSSTLLNPKDSNIEILDAKSSDNITPSFYVNDSDFENV
ST210      NPKSEDSYNTRLTIQGYQEDLNNSSTLLNPKDSNIEILDAKSSDNITPSFYVNDSDFENV
           ********************:*:*******.*::************.***::**.***:*

ST2        TGNFGINQKGDKKAQIDFGHIDHPYIVKVTSKIDPSSSQDLRTRVIMENENAEGTTDFYA
ST57       TGNFGINQKGDKKAQIDFGHIDHPYIVKVTSKIDPSSSQDLRTRVIMENENAEGTTDFYA
ST218      TGNFGINQKGDKKAQIDFGHIDHPYIVKVTSKIDPSSSQDLRTRVIMENENAEGTTDFYA
ST5        TNQYKIDQIGDKKAQIDFGHIDHPYIVKVTSKIDPNSSKDLRTRVIMENENAEGTKDFYV
ST210      TNQYKIDQIGDKKAQIDFGHIDHPYIVKVTSKIDPNSSKDLRTRVIMENENAEGTKDFYV
           *.:: *:* **************************.**:****************.***.

ST2        HDNTVERLGANGVATGNE
ST57       HDNTVERLGANGVATGNE
ST218      HDNTVERLGANGVATGNE
ST5        HDNTVERLGANGVATGNE
ST210      HDNTVERLGANGVATGNE
           ******************
```

Figure 16. Sequence comparison of SesJ A-region from different STs. ST2, ST57, ST218 isolates contain SesJ isoform I. ST5 and ST210 isolates contain SesJ isoform II.

in the *orfX* gene (Figure 15). In *S. epidermidis* NIHLM039 the *sesJ* gene is located on SCC non-*mec* element.

SesJ isoforms

Most MSCRAMMs have isoforms, which are identified based on the sequence difference in the A-region. For example, FnBPA has seven isoforms based on the A-region itself [213]. Within the A-region, most of the differences lie in the N3 region of the MSCRAMMs. Similarly, two different isoforms of SesJ were observed in the thirty *sesJ*$^+$ isolates in the collection. SesJ A-region from the two isoforms, share 94.7% identity with each other, with most differences in the N3 region (Figure 16) [201]. It is worth noting, that SesJ isoforms are present in different clonal complex, independent of SCC*mec* carrying the *sesJ* gene. Both ST5 and ST210 both represent CC5 and encoded SesJ isoform II but ST5 *sesJ*$^+$ isolates harbored SCC*mec* Type IV and ST210 isolates harbored SCC*mec* Type VII. On the other hand, ST2 (CC2) *sesJ*$^+$ isolate contained SesJ isoform I even though these isolates also harbored SCC*mec* Type IV (Table 9).

Increase in SesJ frequency

171 collected isolates were collected from July, 2014 to May, 2016 at MD Anderson hospital. We grouped the isolates into 3 groups based on the date of isolation. 54 isolates used in this study were collected in year 1, 54 isolates in year 2 and 62 isolates in year 3. Frequency of the *sesJ* gene in these isolates rose from 7% in year 1 to 18% in year 2 and 24% in year 3. On the other hand, frequency of the pls gene in year 1 was 9%, and 0% in

ST	CC	*sesJ*$^{+}$/total isolates	Isoform	SCC*mec* Type
ST 2	2	6/34	I	IV
ST 5	5	16/58	II	IV
ST 210	5	6/6	II	VII

Table 9. SesJ isoforms are present in different clonal complexes.

year 2, followed by 6% in year 3 (Figure 17). Thus, the frequency of *sesJ*$^{+}$ *S. epidermidis* isolates has increased three times in three years. This increase was specific to SesJ and not observed for Pls.

Discussion

S. epidermidis has an open pan genome perhaps because of the high frequency of genetic recombination [109]. Accessory genes constitute 20% of the whole genome in *S. epidermidis* [109]. In present work, two of the accessory genes encoding for cell wall anchored proteins, present in SCC*mec* elements were identified. SesJ is present in SCC*mec* Type IV, VII and non-typeable elements. In MSSE, the *sesJ* gene is present in a SCC non-*me*c element containing the *ccr* genes. Although origin of the *sesJ* gene is not known, it's speculated that given the chimeric nature of the protein it was generated through recombination in CoNS. Additionally, we identified the *pls* gene in SCC*mec* Type IV. Pls has been reported before in SCC*mec* Type 1 and ψSCC*pls* [203, 214] in *S. aureus*. CoNS also acts as a reservoir of virulence factors that can be transferred through horizontal gene transfer to other species like *S. aureus* [150]. Presence of the *sesJ* gene in SCC*mec* elements raises the possibility of transfer of the gene to *S. aureus*.

A high degree of genetic diversity was observed in the isolates used in this study. This is consistent with previously reported studies. ST5, ST2 and ST83 accounted for over 60% of the total isolates tested. ST2 is one of the most successful genetic lineages of

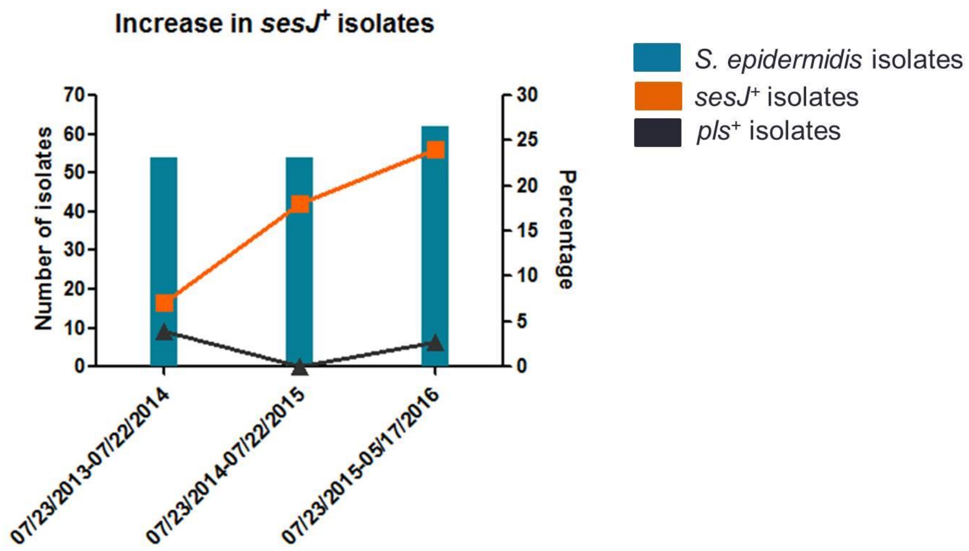


Figure 17. Frequency of the *sesJ*$^{+}$ *S. epidermidis* isolates is increasing.

S. epidermidis spread worldwide [50, 181, 183, 215-217]. Recent reports have documented higher proportion of ST5 in their studies compared to ST2 [50, 217, 218]. Similarly, ST5 clinical isolates predominated in this study. The sequence type ST83 is not as widespread and predominant as ST2 and ST5 but ST83 has been observed in the US, Italy and Greece [50, 136, 219]. We also observed ST210 in our clinical isolates obtained from patients. In another study, ST210 was associated with healthcare personnel only and absent in isolates obtained from patients [215]. Interestingly, all ST210 isolates in our study encode the *sesJ* gene. Perhaps, *S. epidermidis* ST210 isolates can cause infection upon acquiring potential virulence factors like SesJ.

Although this study is limited to one hospital in the US, we reported previously on an epidemiology study for the *sesJ* gene in isolates collected at Columbia University Medical Center [201]. Thus, the *sesJ* gene is not confined to one hospital. In a previous study, we investigated isolates from patients with infected left ventricular assist driveline. This study focused on blood isolates from cancer patients at MD Anderson hospital. We observed a similar percentage of *sesJ*$^{+}$ isolates in both studies. Presence of the *sesJ* gene in clinical isolates from different patients and geographic location shows that *sesJ*$^{+}$ isolates are prevalent in the US.

Both SesJ and Pls are located in multiple SCC*mec* elements. SesJ is the first MSCRAMM encoded on a mobile genetic element. It is also worth noting that (i) SesJ is present in MSSE isolates and (ii) in MSSE isolates, the *sesJ* gene is located in SCC non-

mec element. In our two studies, we observed the *sesJ* gene in commensal isolates although its frequency in commensal isolates from healthy individuals is not known. It was surprising that Pls was only present in MRSE isolates unlike SesJ. Virulence factors like SrdG, *ica* operon, Aap and more, are present in both commensal and infectious isolates. We hypothesize that virulence factors required for establishing infection are also required for the commensal lifestyle of the bacteria.

In this collection, we observed two isoforms of SesJ with 95% identity in the A-region. The A-region is the ligand binding region of the MSCRAMM that evolves through point mutations. Multiple isoforms of other well-studied MSCRAMMs have been identified, with studies underway to understand the differences in the molecular interaction of the isoform with the ligand [213, 220]. The two isoforms of SesJ are present in different clonal complex, which is consistent with *S. aureus* MSCRAMMs (unpublished observation from the lab).

Interestingly, in three years the frequency of SesJ increased 3 fold. This increase was specific to SesJ as the frequency of the *pls* gene in the same collection and similar SCC*mec* element did not increase over the 3 year period. Increase in the *sesJ* gene frequency does not correlate with number of isolates obtained in the respective years or with the number of MRSE isolates alone obtained during the time period. The *sesJ* gene probably utilizes Ccr enzymes to move around. Increase in the frequency of SesJ suggests that SesJ provides a unique advantage to *S. epidermidis*. Pls plays a role in

biofilm formation by utilizing G5 repeats i.e. a similar mechanism as Aap. Unlike *S. aureus*, *S. epidermidis* does not encode as many virulence factors with repetitive or redundant functions. SesJ potentially provides a unique advantage to *S. epidermidis*, which has contributed to its increase in frequency over Pls. Further studies are warranted to uncover the role of SesJ in *S. epidermidis* pathogenesis.

CHAPTER IV

CONCLUSIONS

S. epidermidis does not possess a large arsenal of virulence factors compared to *S. aureus*. Unlike *S. aureus*, *S. epidermidis* causes low grade but persistent infections that are hard to treat. Most clinical isolates of *S. epidermidis* carry multi drug resistance, which limits the antibiotics available to the patient. Biofilm matrix produced by *S. epidermidis* isolates creates a barrier for the antibiotics to target the embedded cells. In addition, antibiotics target live bacterial cells whereas biofilm also contains metabolically inactive bacterial cells that are resistant to antibiotics. Thus, these features make it challenging to treat *S. epidermidis* infections. In the past two decades, the identity and mechanism of action of some of the *S. epidermidis* virulence factors have been discovered but there is still a big gap in understanding how *S. epidermidis* causes infections. Furthermore, most of the characterized virulence factors are encoded in the core genome, whereas 20% of the *S. epidermidis* genome consists of variable genes which can alter the pathogenic potential of the isolate [109]. Therefore, a better understanding of the *S. epidermidis* virulence factors and their role in establishing infection will help identify potential drug targets. The development of new drugs that function by blocking the interaction of the virulence factor with the host ligand provides an alternative to conventional antibiotics.

Studies described in this thesis revealed two, previously unidentified, potential virulence

factors of *S. epidermidis* encoded on a MGE: SesJ and Pls. SesJ is a chimeric cell wall anchored protein that is expressed on the surface of *S. epidermidis*. SesJ is also a prototype of a novel NTR containing subfamily of MSCRAMM. SesJ contains conserved features of an MSCRAMM as well as regions like the NTR and ACR, which have not been found in MSCRAMM before. The NTR-region is present after the signal sequence and contains a repeated 15 amino acid motif. There are the only two other *S. epidermidis* CWA proteins known to contain NTRs i.e. Aap and Pls. The A-region of SesJ has characteristic features found in MSCRAMMs; it contains a two tandem IgG-like fold containing the trench motif and the latch region. MSCRAMMs bind to their ligands using the DLL mechanism. The presence of unique characteristics of MSCRAMM in SesJ indicates that the protein potentially binds to its ligand using the DLL mechanism. SesJ also contains two B-repeats similar to the B-repeats in the Sdr sub-family of MSCRAMMS. In addition, the ACR region is present near the C-terminus and contains 20 amino acid long repeat motif similar to serine rich region of SRRPs from *Streptococcus*. At the C-terminus SesJ contains LPXTG motif, transmembrane domain and a small cytoplasmic tail similar to other known MSCRAMMs.

Three structural homologs of SesJ were also identified in other coagulase negative staphylococci species. SesJ structural homologs ScsJ, ShsJ and SdrI proteins are present in *S. capitis*, *S. haemolyticus* and *S. saprophyticus,* respectively. SesJ and its structural homologs have identical domain organization as well as characteristics unique to MSCRAMMS. Like SesJ, the structural homologs also contain N-terminal signal

sequence, the NTR region, A-region, two B-repeats each, ACR region, LPXTG motif, transmembrane domain and a short cytoplasmic tail. Similar to the SesJ protein, N2N3 region of ScsJ, ShsJ and SdrI proteins also comprise of two linked IgG-like fold containing a predicted latch sequence and the trench motif. Together these proteins form the NTR containing sub-family of MSCRAMMs present in coagulase negative staphylococci.

Although SesJ, ScsJ, ShsJ and SdrI are structurally similar, the A-region of SesJ and its structural homologs have amino acid sequence identities as low as 30%, indicating that these proteins most likely bind to different ligands. For example, SdrI has been shown to bind to fibronectin while my preliminary experiments revealed that SesJ does not interact with fibronectin. The function of SesJ, ScsJ and ShsJ is not known. Along with the MSCRAMM features, the NTR regions of all these proteins also share a conserved motif. Even though the function of the NTRs is not known, the conserved motif suggests a shared function amongst these proteins.

Next, epidemiology studies revealed that the *sesJ* gene is present in ~18% of the clinical isolates tested. In these studies, blood isolates from patients with infected left ventricular assist devices, cancer patients with bloodstream infections, and colonizing isolates from HIV infected patients were used. The *sesJ* gene is present in all the different types of clinical isolates tested, suggesting a broader potential role in the bacterial pathogenesis. Unlike *S. aureus* that is only found in 30% of the population, *S. epidermidis* is

ubiquitous on human skin. It is therefore not surprising that *S. epidermidis* encodes factors required in both commensal and pathogenic lifestyle. These virulence factors could help the bacteria maintain a commensal lifestyle but when given an opportunity also act as virulence factors required in infection [28]. This is evident with the *sdrG*, *sdrF*, *icaADBC*, *aap* genes which are present in both commensal and bloodstream isolates. Likewise, the *sesJ* gene is present in both commensal and bloodstream isolates.

Additionally, the *sesJ* gene is encoded on a composite SCC*mec* Type IV, Type VII and non typeable elements. In methicillin sensitive isolates, the *sesJ* gene is present in a ψSCC element containing the *ccr* genes but missing the *mec* gene complex. SesJ is the first MSCRAMM known to be encoded on a mobile genetic element. All other known MSCRAMMs belonging to Sdr-Clf, Fnbs and Cna family are encoded within the core genome. The presence of the *sesJ* gene in a MGE explains the low frequency of the gene in clinical isolates compared to other staphylococcal MSCRAMMS. The methicillin resistance gene is acquired by the bacteria under antibiotic pressure. The *sesJ* gene most likely hitchhikes with the *mec* and *ccr* gene complex and is acquired by the neighboring bacteria. Interestingly, the frequency of *sesJ*$^{+}$ isolates has increased 3 fold in only 3 years at MD Anderson. Antibiotic pressure and presence of the *sesJ* gene in a MGE has most likely contributed to its increase in frequency.

Pls, another cell wall anchored protein, was identified in a SCC*mec* Type IV in this study. At the N-terminal of the Pls protein is the signal sequence, followed by NTR-

region (similar to that found in SesJ), the A-region, G5 repeats (involved in biofilm formation), the SD repeat (as found in Sdr MSCRAMM) and the cell wall anchoring region [204]. Previously, the Pls protein has been shown to be associated with SCC*mec* Type I in *S. aureus*, *S. capitis*, *S. haemolyticus* and *S. epidermidis*, and ψSCC element in *S. aureus* WA MRSA-40 [214]. As suggested by the multiple different types, sub-types and multiple novel SCC elements reported in the literature, SCC*mec* elements are a hotspot for recombination [138, 192-195, 221]. It is perhaps the result of recombination that the *pls* gene is also present in SCC*mec* Type IV element. In *S. aureus*, Pls protein has been reported to play a role in septic arthritis as well as biofilm formation. However, it's role in *S. epidermidis* pathogenesis has not been studied. It can be assumed that Pls plays a role in biofilm formation in *S. epidermidis* too.

Although, a function of SesJ is not known yet, studies are underway to identify its role in *S. epidermidis* pathogenesis. Compared to *S. epidermidis*, *S. aureus* MSCRAMMs are well studied and functions are known for most of the MSCRAMMs. Based on the previous studies, MSCRAMMs have been shown to play a role in adhesion of the bacteria to the host tissue, mediating invasion of the bacteria into host cells, evading immune response and biofilm formation. To this end, I tested binding of the SesJ protein to known MSCRAMMs ligands including plasma proteins (fibrinogen from mouse, rat, rabbit, pig, sheep, bovine, dog and human, fibronectin), extracellular matrix proteins (vitronectin, laminin, fibromodulin, cytokeratin 10, decorin core, decorin, biglycan core, biglycan, collagen I/II/III/IV) as well as complement proteins (factor I, factor H, factor

B, factor D, c1q). SdrG interaction with human protein fibrinogen was used as a positive control for the adhesion assays. First, binding of the recombinant SesJ A-region was tested in a solid phase assay where potential ligands were immobilized on a polystyrene plate. Next, bacterial attachment assays were done utilizing full length SesJ protein expressed on the surface of *Lactococcus lactis,* a heterologous host. Similar to solid phase assay, host proteins were immobilized on a polystyrene plate and incubated with the *L. lacis* expressing SdrG or SesJ. *L. lactis* containing empty vector was used as negative control to rule out interaction of *L. lactis* CWA proteins with the tested host proteins. While, positive control SdrG interaction could be observed in these assays, SesJ did not interact with any of the proteins tested.

Most MSCRAMMs bind to plasma proteins e.g. ClfA/B, FnbpA/B, Bbp bind to fibrinogen present in the plasma, FnbpA/B bind to fibronectin. To avoid immobilizing the host proteins on a polystyrene plate, pull down assays were developed to test binding of the soluble recombinant SesJ protein with the soluble plasma/serum proteins. First, histidine tag system was established for the pull down experiment. However, high background was observed in pull down assays with histidine tagged recombinant protein. Therefore, Strep-tagged recombinant SesJ AB region was purified and streptactin resin was used to capture the strep-tagged protein. Although interaction of the strep tagged FnbpA protein could be observed with fibrinogen and fibronectin in the plasma, no interactions were observed for the SesJ protein. These experiments show SesJ did not bind to common MSCRAMM ligands like fibrinogen and fibronectin.

Although a ligand for recombinant SesJ AB could not be identified with the pull down assay, interaction of SesJ with proteins present in low concentration in plasma cannot be ruled out. Pull down assays can identify stable or strong interacting ligands present in sufficient concentration to show a band on the SDS page gel. However, weak or transient interactions, and strong interactions with proteins present in low concentration are often missed by the pull down assays. Perhaps a broader ligand screening assays like phage display can help identify a ligand for the SesJ protein. Phage display system was used to identify a ligand for MSCRAMM SdrC, which provides a proof of concept that this technology can be used for identifying MSCRAMM ligands [104]. Additionally, phage display system does not depend on the concentration of the ligand but is limited by the type of library used for the assay, and has a higher false positive rate. On the other hand, it can be hypothesized that the glycosylation of the SesJ protein is required for its binding to the ligand. To test this hypothesis, the *gftA* and *gtfB* gene, encoded next to the *sesJ* gene in SCC*mec* cassette can be transformed into the *L. lactis* expressing full length SesJ protein on the surface. The new constructs can be used in a bacterial attachment assays.

MSCRAMMs have also been shown to play a role in immune evasion. Cna, *S. aureus* MSCRAMM inhibits classical complement pathway by binding to the collagen like stalk of C1q [103]. Similarly Bbp and SdrE proteins interact with Factor H and inhibit alternative complement pathway [107]. To test the potential role of SesJ in immune evasion, classical complement pathway inhibition as well as phagocytosis inhibition was

tested. Inhibition of classical complement pathway was tested by monitoring complement mediated hemolysis of antibody sensitized sheep erythrocytes. MSCRAMM Cna was used as positive control which has been demonstrated to inhibit classical complement pathway. Recombinant Cna and SesJ A-region was incubated with normal human serum before addition of the antibody sensitized sheep erythrocytes. The inhibition of hemolysis was monitored by measuring the absorbance of the supernatant. While, Cna did inhibit hemolysis of sheep erythrocyte, no effect was observed for SesJ.

Phagocytosis inhibition by the SesJ protein was tested using freshly isolated human neutrophils from a healthy individual. A secreted *S. aureus* protein Efb that inhibits phagocytosis was used as a positive control. FITC labeled bacteria was mixed with either normal human serum or normal human plasma in the presence of recombinant Efb or SesJ AB. Next, freshly isolated neutrophils were added to the mixture. Phagocytosis was detected using flow cytometry. While Efb demonstrated phagocytosis inhibition as expected, SesJ did not cause inhibition. While SesJ did not stimulate an immune evasion in the assays tested, possible role of SesJ in immune evasion cannot be completely ruled out. Biofilm formation is one of the mechanisms used by *S. epidermidis* to immune evasion. SesJ role in biofilm formation is currently under investigation. Also, *S. epidermidis* causes low grade and persistent infection and most likely evades immune response through multiple mechanisms. Moreover, *S. aureus* has numerous mechanisms for immune evasion that have not been yet studied in *S. epidermidis*. Thus, detailed studies are required to discover if and how SesJ plays a role in immune evasion.

The role of SesJ in biofilm formation is currently under investigation. In *S. aureus*, MSSA isolates form primarily polysaccharide dependent biofilm whereas MRSA isolates form protein dependent biofilm [222]. The quorum sensing system *agr* is regulated differentially in MRSA and MSSA isolates. The *agr* operon is upregulated in MSSA isolates which in turn upregulates PIA and secreted proteases production. Secreted proteases cleave the surface protein, resulting in a biofilm formation composed of PIA. On the other hand, in MRSA isolates, the *agr* operon is downregulated, which inhibits PIA production and upregulates surface protein expression. Therefore, MRSA isolates form a surface protein dependent biofilm formation. The presence of the *sesJ* gene in SCC*mec* element along with similarities of NTR region of SesJ to biofilm forming protein Aap and Pls point to the potential role of SesJ in biofilm formation [180, 204]. Additionally, MSCRAMMs like FnbpA and SdrC promote biofilm formation through homophilic interactions of the N2N3 region. Blood isolates obtained from cancer patients presented with blood stream infections will be used for the static biofilm formation assays, followed up by biofilm experiments performed under flow to mimic *in vivo* conditions.

These studies have revealed that SesJ has characteristics of a virulence factor and its frequency is increasing at MD Anderson. In my study, I observed a 3 fold increase in the frequency of the *sesJ* gene over the 3 year period. The observed increase was specific to the SesJ protein and not observed for the Pls protein. This increase did not correlate to the total number of isolates collected in a year, number of methicillin resistant isolates

collected or to clonal expansion. Increase in the *sesJ* gene frequency from year 1 – year 2 was observed because of an increase in the *sesJ*$^{+}$ ST 5B isolates whereas, the total number of ST5B isolates collected in year 1 and year 2 was similar. Increase in year 2 – year 3 was caused by an increase of the *sesJ*$^{+}$ ST210 isolates. Increase in the *sesJ*$^{+}$ isolates at MD Anderson emphasizes the need to discover a ligand for SesJ and its role in pathogenesis and will be the focus of future studies. Different isoforms of SesJ may bind to the ligand with different affinities and modulate the effect of the interaction in *S. epidermidis* pathogenesis.

It is known that the expression of the *mecA* operon is increased in the presence of the methicillin. Also, MRSA isolates form a protein dependent biofilm. While SesJ is constitutively expressed on the surface of *S. epidermidis* isolate tested, it could be the basal level of SesJ expression. It is likely that expression of the SesJ protein increases in the presence of methicillin and be regulated by a molecule within the SCC*mec* cassette. Presence of the SesJ protein in SCC*mec* cassette in a MRSE isolate indicates that SesJ plays a role in protein dependent biofilm formation in MRSE isolates. It is plausible that the NTR region of SesJ protein plays a role in the attachment phase of the biofilm formation and the A-region promotes accumulation phase through homophilic interactions. These possible mechanisms will be tested in the biofilm formation assays.

It is worth noting that the SCC*mec* cassette encodes for Ccr enzymes responsible for the insertion and excision of the cassette. Some future studies will also be directed to

understand if *sesJ* and associated genes move around independently or by hijacking the Ccr enzymes. In addition, the analysis of clinical outcome data from cancer patients will help understand whether the presence or absence of the *sesJ* /*pls* gene correlate with clinical outcome in patients. From my analysis I can conclude that the *sesJ* gene is present in multiple STs and SCC*mec* elements. Clinical outcome data analysis would help understand if the presence or absence of the *sesJ* gene in a particular ST or SCC*mec* correlates with clinical outcome parameters like polymicrobial infections, persistent disease, high CFUs obtained or overall severity of the disease.

Overall, work in this thesis expands our understanding on potential virulence factors present in the SCC*mec* elements in *S. epidermidis* as well as the MSCRAMM family of proteins. Cell wall anchored proteins are critical virulence factors for all Gram-positive pathogens. The MSCRAMMs, a family of cell wall anchored proteins, include important virulence factors present mostly in *Staphylococcus*. Based on this work, the MSCRAMM family of proteins has expanded from three sub-families to four. It also revealed SesJ is the first MSCRAMM present in MGE. Additionally, most *S. epidermidis* studies have been done with isolates from patients with an infected medical device. There is considerably less information on clinical *S. epidermidis* isolates from cancer patients. Work in this thesis utilized blood isolates from cancer patients with blood stream infections and provided the frequency of previously known and new cell wall anchored proteins as well as SCC*mec* in these isolates. Personalized drugs are the new future of medicine. Although personalized medicine has been beneficial in the field

of oncology, it has not been expanded to treat difficult bacterial infections. A better understanding of *S. epidermidis* virulence factors and their interactions with specific host targets may help create drugs specific to the clinical isolate causing an infection.

REFERENCES

[1] Magill, S.S., Edwards, J.R., Bamberg, W., Beldavs, Z.G., Dumyati, G., et al., Multistate point-prevalence survey of health care-associated infections. N Engl J Med, 2014. **370**(13): p. 1198-1208.

[2] Becker, K., Heilmann, C., and Peters, G., Coagulase-negative staphylococci. Clin Microbiol Rev, 2014. **27**(4): p. 870-926.

[3] Grice, E.A. and Segre, J.A., The skin microbiome. Nat Rev Microbiol, 2011. **9**(4): p. 244-253.

[4] Kloos, W.E. and Musselwhite, M.S., Distribution and persistence of *Staphylococcus* and *Micrococcus* species and other aerobic bacteria on human skin. Appl Microbiol, 1975. **30**(3): p. 381-385.

[5] Kloos, W.E. and Schleifer, K.H., Isolation and characterization of staphylococci from human skin II. Descriptions of Four new species: *Staphylococcus warneri, Staphylococcus capitis, Staphylococcus hominis, and Staphylococcus simulans.* Int J Syst Evol Microbiol, 1975. **25**(1): p. 62-79.

[6] Schleifer, K.H. and Kloos, W.E., Isolation and characterization of staphylococci from human skin I. Amended descriptions of *Staphylococcus epidermidis* and *Staphylococcus saprophyticus* and descriptions of three new species: *Staphylococcus cohnii, Staphylococcus haemolyticus,* and *Staphylococcus xylosus.* Int J Syst Evol Microbiology, 1975. **25**(1): p. 50-61.

[7] Weiner, L.M., Webb, A.K., Limbago, B., Dudeck, M.A., Patel, J., et al.,

Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention*, 2011-2014.* Infect Control Hosp Epidemiol, 2016. **37**(11): p. 1288-1301.

[8] Warren, D.K., Quadir, W.W., Hollenbeak, C.S., Elward, A.M., Cox, M.J., et al., Attributable cost of catheter-associated bloodstream infections among intensive care patients in a nonteaching hospital. Crit Care Med, 2006. **34**(8): p. 2084-2089.

[9] Blot, S.I., Depuydt, P., Annemans, L., Benoit, D., Hoste, E., et al., Clinical and economic outcomes in critically ill patients with nosocomial catheter-related bloodstream infections. Clin Infect Dis, 2005. **41**(11): p. 1591-1598.

[10] Rello, J., Ochagavia, A., Sabanes, E., Roque, M., Mariscal, D., et al., Evaluation of outcome of intravenous catheter-related infections in critically ill patients. Am J Respir Crit Care Med, 2000. **162**(3 Pt 1): p. 1027-1030.

[11] Zimlichman, E., Henderson, D., Tamir, O., Franz, C., Song, P., et al., Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system. JAMA Intern Med, 2013. **173**(22): p. 2039-2046.

[12] Chu, V.H., Woods, C.W., Miro, J.M., Hoen, B., Cabell, C.H., et al., Emergence of coagulase-negative staphylococci as a cause of native valve endocarditis. Clin Infect Dis, 2008. **46**(2): p. 232-242.

[13] Simonsen, K.A., Anderson-Berry, A.L., Delair, S.F., and Davies, H.D., Early-onset neonatal sepsis. Clin Microbiol Rev, 2014. **27**(1): p. 21-47.

[14] Didier, C., Streicher, M.P., Chognot, D., Campagni, R., Schnebelen, A., et al., Late-onset neonatal infections: incidences and pathogens in the era of antenatal antibiotics. Eur J Pediatr, 2012. **171**(4): p. 681-687.

[15] Coullioud, D., Van der Auwera, P., Viot, M., and Lasset, C., Prospective multicentric study of the etiology of 1051 bacteremic episodes in 782 cancer patients. CEMIC (French-Belgian Study Club of Infectious Diseases in Cancer). Support Care Cancer, 1993. **1**(1): p. 34-46.

[16] Shin, J.H., Kim, S.H., Jeong, H.S., Oh, S.H., Kim, H.R., et al., Identification of coagulase-negative staphylococci isolated from continuous ambulatory peritoneal dialysis fluid using 16S ribosomal RNA, tuf, and SodA gene sequencing. Perit Dial Int, 2011. **31**(3): p. 340-346.

[17] Petinaki, E., Kontos, F., Miriagou, V., Maniati, M., Hatzi, F., et al., Survey of methicillin-resistant coagulase-negative staphylococci in the hospitals of central Greece. Int J Antimicrob Agents, 2001. **18**(6): p. 563-566.

[18] Takeuchi, F., Watanabe, S., Baba, T., Yuzawa, H., Ito, T., et al., Whole-genome sequencing of *Staphylococcus haemolyticus* uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. J Bacteriol, 2005. **187**(21): p. 7292-7308.

[19] Latham, R.H., Running, K., and Stamm, W.E., Urinary tract infections in young adult women caused by *Staphylococcus saprophyticus.* JAMA, 1983. **250**(22): p. 3063-3066.

[20] Rasigade, J.P., Raulin, O., Picaud, J.C., Tellini, C., Bes, M., et al., Methicillin-

resistant *Staphylococcus capitis* with reduced vancomycin susceptibility causes late-onset sepsis in intensive care neonates. PLoS One, 2012. **7**(2): p. e31548. doi: 10.1371/journal.pone.0031548

[21] Yao, Y., Sturdevant, D.E., and Otto, M., Genomewide analysis of gene expression in *Staphylococcus epidermidis* biofilms: insights into the pathophysiology of *S. epidermidis* biofilms and the role of phenol-soluble modulins in formation of biofilms. J Infect Dis, 2005. **191**(2): p. 289-298.

[22] Resch, A., Rosenstein, R., Nerz, C., and Gotz, F., Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. Appl Environ Microbiol, 2005. **71**(5): p. 2663-2676.

[23] Fey, P.D. and Olson, M.E., Current concepts in biofilm formation of *Staphylococcus epidermidis.* Future Microbiol, 2010. **5**(6): p. 917-933.

[24] Otto, M., Molecular basis of *Staphylococcus epidermidis* infections. Semin Immunopathol, 2012. **34**(2): p. 201-214.

[25] Otto, M., *Staphylococcus epidermidis*-the 'accidental' pathogen. Nat Rev Microbiol, 2009. **7**(8): p. 555-567.

[26] Heilmann, C., Hussain, M., Peters, G., and Götz, F., Evidence for autolysin-mediated primary attachment of Staphylococcus epidermidis to a polystyrene surface. Mol Microbiol, 1997. **24**(5): p. 1013-1024.

[27] Heilmann, C., Thumm, G., Chhatwal, G.S., Hartleib, J., Uekotter, A., et al., Identification and characterization of a novel autolysin (Aae) with adhesive properties from *Staphylococcus epidermidis.* Microbiology, 2003. **149**(Pt 10): p.

2769-2778.

[28] Herrmann, M., Vaudaux, P.E., Pittet, D., Auckenthaler, R., Lew, P.D., et al., Fibronectin, fibrinogen, and laminin act as mediators of adherence of clinical staphylococcal isolates to foreign material. J Infect Dis, 1988. **158**(4): p. 693-701.

[29] Foster, T.J., Geoghegan, J.A., Ganesh, V.K., and Hook, M., Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus.* Nat Rev Microbiol, 2014. **12**(1): p. 49-62.

[30] Davis, S.L., Gurusiddappa, S., McCrea, K.W., Perkins, S., and Höök, M., *SdrG, a* fibrinogen-binding bacterial adhesin of the microbial surface components recognizing adhesive matrix molecules subfamily from *Staphylococcus epidermidis,* targets the thrombin cleavage site in the Bβ chain. J Biol Chem, 2001. **276**(30): p. 27799-27805.

[31] Pei, L. and Flock, J.I., Lack of fbe, the gene for a fibrinogen-binding protein from *Staphylococcus epidermidis*, reduces its adherence to fibrinogen coated surfaces. Microb Pathog, 2001. **31**(4): p. 185-193.

[32] Arrecubieta, C., Lee, M.-H., Macey, A., Foster, T.J., and Lowy, F.D., SdrF, a *Staphylococcus epidermidis* surface protein, binds Type I collagen. J Biol Chem, 2007. **282**(26): p. 18767-18776.

[33] Arrecubieta, C., Asai, T., Bayern, M., Loughman, A., Fitzgerald, J.R., et al., The role of *Staphylococcus aureus* adhesins in the pathogenesis of ventricular assist device-related infections. J Infect Dis, 2006. **193**(8): p. 1109-1119.

[34] Williams, R.J., Henderson, B., Sharp, L.J., and Nair, S.P., Identification of a fibronectin-binding protein from *Staphylococcus epidermidis.* Infect Immun, 2002. **70**(12): p. 6805-6810.

[35] Christner, M., Franke, G.C., Schommer, N.N., Wendt, U., Wegert, K., et al., The giant extracellular matrix-binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment to fibronectin. Mo Microbiol, 2010. **75**(1): p. 187-207.

[36] Qin, Z., Ou, Y., Yang, L., Zhu, Y., Tolker-Nielsen, T., et al., Role of autolysin-mediated DNA release in biofilm formation of *Staphylococcus epidermidis.* Microbiology, 2007. **153**(Pt 7): p. 2083-2092.

[37] Mann, E.E., Rice, K.C., Boles, B.R., Endres, J.L., Ranjit, D., et al., Modulation of eDNA release and degradation affects *Staphylococcus aureus* biofilm maturation. PLoS One, 2009. **4**(6): p. e5822. doi: 10.1371/journal.pone.0005822

[38] Mack, D., Fischer, W., Krokotsch, A., Leopold, K., Hartmann, R., et al., The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: purification and structural analysis. J Bacteriol, 1996. **178**(1): p. 175-183.

[39] Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., et al., Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis.* Mol Microbiol, 1996. **20**(5): p. 1083-1091.

[40] Rohde, H., Burandt, E.C., Siemssen, N., Frommelt, L., Burdelski, C., et al., Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of

Staphylococcus epidermidis and *Staphylococcus aureus* isolated from prosthetic hip and knee joint infections. Biomaterials, 2007. **28**(9): p. 1711-1720.

[41] Rohde, H., Burdelski, C., Bartscht, K., Hussain, M., Buck, F., et al., Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. Mol Microbiol, 2005. **55**(6): p. 1883-1895.

[42] Conrady, D.G., Brescia, C.C., Horii, K., Weiss, A.A., Hassett, D.J., et al., A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19456-19461.

[43] Tormo, M.A., Knecht, E., Gotz, F., Lasa, I., and Penades, J.R., Bap-dependent biofilm formation by pathogenic species of *Staphylococcus*: evidence of horizontal gene transfer? Microbiology, 2005. **151**(Pt 7): p. 2465-2475.

[44] Lister, J.L. and Horswill, A.R., *Staphylococcus aureus* biofilms: recent developments in biofilm dispersal. Front Cell Infect Microbiol, 2014. **4**: p. 178-187. doi: 10.3389/fcimb.2014.00178

[45] Vuong, C., Gerke, C., Somerville, G.A., Fischer, E.R., and Otto, M., Quorum-sensing control of biofilm factors in *Staphylococcus epidermidis.* J Infect Dis, 2003. **188**(5): p. 706-718.

[46] Kong, K.F., Vuong, C., and Otto, M., *Staphylococcus* quorum sensing in biofilm formation and infection. Int J Med Microbiol, 2006. **296**(2-3): p. 133-139.

[47] Wang, R., Khan, B.A., Cheung, G.Y., Bach, T.H., Jameson-Lee, M., et al.,

Staphylococcus epidermidis surfactant peptides promote biofilm maturation and dissemination of biofilm-associated infection in mice. J Clin Invest, 2011. **121**(1): p. 238-248.

[48] Boles, B.R. and Horswill, A.R., Agr-mediated dispersal of *Staphylococcus aureus* biofilms. PLoS Pathog, 2008. **4**(4): p. e1000052.

[49] Gordon, R.J., Miragaia, M., Weinberg, A.D., Lee, C.J., Rolo, J., et al., *Staphylococcus epidermidis* colonization is highly clonal across US cardiac centers. J Infect Dis, 2012. **205**(9): p. 1391-1398.

[50] Mendes, R.E., Deshpande, L.M., Costello, A.J., and Farrell, D.J., Molecular epidemiology of *Staphylococcus epidermidis* clinical isolates from U.S. hospitals. Antimicrob Agents Chemother, 2012. **56**(9): p. 4656-4661.

[51] Arrecubieta, C., Toba, F.A., von Bayern, M., Akashi, H., Deng, M.C., et al., SdrF, a *Staphylococcus epidermidis* surface protein, contributes to the initiation of ventricular assist device driveline-related infections. PLoS Pathog, 2009. **5**(5): p. e1000411. doi: 10.1371/journal.ppat.1000411

[52] Park, P.W., Rosenbloom, J., Abrams, W.R., Rosenbloom, J., and Mecham, R.P., Molecular cloning and expression of the gene for Elastin-binding Protein (ebpS) in *Staphylococcus aureus.* J Biol Chem, 1996. **271**(26): p. 15803-15809.

[53] Zhang, Y.-Q., Ren, S.-X., Li, H.-L., Wang, Y.-X., Fu, G., et al., Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). Mol Microbiol, 2003. **49**(6): p. 1577-1593.

[54] Spiliopoulou, A.I., Krevvata, M.I., Kolonitsiou, F., Harris, L.G., Wilkinson, T.S.,

et al., An extracellular *Staphylococcus epidermidis* polysaccharide: relation to Polysaccharide Intercellular Adhesin and its implication in phagocytosis. BMC Microbiol, 2012. **12**(1): p. 76-90. doi: 10.1186/1471-2180-12-76

[55] Arciola, C.R., Baldassarri, L., and Montanaro, L., In catheter infections by *Staphylococcus epidermid*is the intercellular adhesion (*ica*) locus is a molecular marker of the virulent slime-producing strains. J Biomed Mater Res, 2002. **59**(3): p. 557-562.

[56] Rohde, H., Burdelski, C., Bartscht, K., Hussain, M., Buck, F., et al., Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. Mol Microbiol, 2005. **55**(6): p.1883-1895.

[57] Macintosh, R.L., Brittan, J.L., Bhattacharya, R., Jenkinson, H.F., Derrick, J., et al., The terminal A domain of the fibrillar Accumulation-Associated Protein (Aap) of *Staphylococcus epidermidis* mediates adhesion to human corneocytes. J Bacteriol, 2009. **191**(22): p. 7007-7016.

[58] Sadovskaya, I., Vinogradov, E., Flahaut, S., Kogan, G., and Jabbouri, S., Extracellular carbohydrate-containing polymers of a model biofilm-producing strain, *Staphylococcus epidermidis RP62A.* Infect and Immun, 2005. **73**(5): p. 3007-3017.

[59] Kocianova, S., Vuong, C., Yao, Y., Voyich, J.M., Fischer, E.R., et al., Key role of poly-γ-dl-glutamic acid in immune evasion and virulence of *Staphylococcus epidermidis.* J Clin Invest, 2005. **115**(3): p. 688-694.

[60] Teufel, P. and Gotz, F., Characterization of an extracellular metalloprotease with elastase activity from *Staphylococcus epidermidis.* J Bacteriol, 1993. **175**(13): p. 4218-4224.

[61] Lai, Y., Villaruz, A.E., Li, M., Cha, D.J., Sturdevant, D.E., et al., The human anionic antimicrobial peptide dermcidin induces proteolytic defence mechanisms in staphylococci. Mol Microbiol, 2007. **63**(2): p. 497-506.

[62] Li, M., Lai, Y., Villaruz, A.E., Cha, D.J., Sturdevant, D.E., et al., Gram-positive three-component antimicrobial peptide-sensing system. Proc Natl Acad Sci U S A, 2007. **104**(22): p. 9469-9474.

[63] Cheung, G.Y.C., Rigby, K., Wang, R., Queck, S.Y., Braughton, K.R., et al., *Staphylococcus epidermidis* strategies to avoid killing by human neutrophils. PLOS Pathog, 2010. **6**(10): p. e1001133. doi: 10.1371/journal.ppat.1001133

[64] Mehlin, C., Headley, C.M., and Klebanoff, S.J., An inflammatory polypeptide complex from *Staphylococcus epidermidis*: isolation and characterization. Journal of Exp Med, 1999. **189**(6): p. 907-918.

[65] Liles, W.C., Thomsen, A.R., O'Mahony, D.S., and Klebanoff, S.J., Stimulation of human neutrophils and monocytes by staphylococcal phenol-soluble modulin. J Leukoc Biol, 2001. **70**(1): p. 96-102.

[66] Kaito, C., Saito, Y., Nagano, G., Ikuo, M., Omae, Y., et al., Transcription and translation products of the cytolysin gene *psm-mec* on the mobile genetic element SCC*mec* regulate *Staphylococcus aureus* virulence. PLOS Pathog, 2011. **7**(2): p. e1001267. doi: 10.1371/journal.ppat.1001267

[67] Madhusoodanan, J., Seo, K.S., Remortel, B., Park, J.Y., Hwang, S.Y., et al., An enterotoxin-bearing pathogenicity island in *Staphylococcus epidermidis.* J Bacteriol, 2011. **193**(8): p. 1854-1862.

[68] Chamberlain, N.R. and Brueggemann, S.A., Characterisation and expression of fatty acid modifying enzyme produced by *Staphylococcus epidermidis.* J Med Microbiol, 1997. **46**(8): p. 693-697.

[69] Khodaparast, L., Khodaparast, L., Shahrooei, M., Stijlemans, B., Merckx, R., et al., The possible role of *Staphylococcus epidermidis* LPxTG surface protein SesC in biofilm formation. PLoS One, 2016. **11**(1): p. e0146704. doi: 10.1371/journal.pone.0146704

[70] Shahrooei, M., Hira, V., Khodaparast, L., Khodaparast, L., Stijlemans, B., et al., Vaccination with SesC decreases *Staphylococcus epidermidis* biofilm formation. Infect Immun, 2012. **80**(10): p. 3660-3668.

[71] Lizcano, A., Sanchez, C.J., and Orihuela, C.J., A role for glycosylated serine-rich repeat proteins in gram-positive bacterial pathogenesis. Mol Oral Microbiol, 2012. **27**(4): p. 257-269.

[72] Patti, J.M., Allen, B.L., McGavin M.J., and Hook M., MSCRAMM-mediated adherence of microorganisms to host tissue. Annu Rev microbiol, 1994. **48**: p.585-617.

[73] Spirig, T., Weiner, E.M., and Clubb, R.T., Sortase enzymes in Gram-positive bacteria. Mol Microbiol, 2011. **82**(5): p. 1044-1059.

[74] Marraffini, L.A., Dedent, A.C., and Schneewind, O., Sortases and the art of

anchoring proteins to the envelopes of gram-positive bacteria. Microbiol Mol Biol Rev, 2006. **70**(1): p. 192-221.

[75] Navarre, W.W. and Schneewind, O., Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Mol Biol Rev, 1999. **63**(1): p. 174-229.

[76] Bowden, M.G., Chen, W., Singvall, J., Xu, Y., Peacock, S.J., et al., Identification and preliminary characterization of cell-wall-anchored proteins of *Staphylococcus epidermidis.* Microbiology, 2005. **151**(Pt 5): p. 1453-1464.

[77] Foster, T.J. and Hook, M., Surface protein adhesins of *Staphylococcus aureus.* Trends Microbiol, 1998. **6**(12): p. 484-488.

[78] Hartford, O., Francois, P., Vaudaux, P., and Foster, T.J., The dipeptide repeat region of the fibrinogen-binding protein (clumping factor) is required for functional expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. Mol Microbiol, 1997. **25**(6): p. 1065-1076.

[79] McCrea, K.W., Hartford, O., Davis, S., Eidhin, D.N., Lina, G., et al., The serine-aspartate repeat (Sdr) protein family in *Staphylococcus epidermidis.* Microbiology, 2000. **146 (Pt 7)**: p. 1535-1546.

[80] Brouillette, E., Grondin, G., Shkreta, L., Lacasse, P., and Talbot, B.G., In vivo and in vitro demonstration that S*taphylococcus aureus* is an intracellular pathogen in the presence or absence of fibronectin-binding proteins. Microb Pathog, 2003. **35**(4): p. 159-168.

[81] Peacock, S.J., Foster, T.J., Cameron, B.J., and Berendt, A.R., Bacterial

fibronectin-binding proteins and endothelial cell surface fibronectin mediate adherence of *Staphylococcus aureus* to resting human endothelial cells. Microbiology, 1999. **145** (Pt 12): p. 3477-3486.

[82] Que, Y.A., Haefliger, J.A., Piroth, L., Francois, P., Widmer, E., et al., Fibrinogen and fibronectin binding cooperate for valve infection and invasion in *Staphylococcus aureus* experimental endocarditis. J Exp Med, 2005. **201**(10): p. 1627-1635.

[83] Bingham, R.J., Rudino-Pinera, E., Meenan, N.A., Schwarz-Linek, U., Turkenburg, J.P., et al., Crystal structures of fibronectin-binding sites from *Staphylococcus aureus* FnBPA in complex with fibronectin domains. Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12254-12258.

[84] Deivanayagam, C.C., Wann, E.R., Chen, W., Carson, M., Rajashankar, K.R., et al., A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. EMBO J, 2002. **21**(24): p. 6660-6672.

[85] Zong, Y., Xu, Y., Liang, X., Keene, D.R., Hook, A., et al., A Collagen Hug model for *Staphylococcus aureus* CNA binding to collagen. EMBO J, 2005. **24**(24): p. 4224-4236.

[86] Deivanayagam, C.C., Rich, R.L., Carson, M., Owens, R.T., Danthuluri, S., et al., Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. Structure, 2000. **8**(1): p. 67-78.

[87] Ponnuraj, K., Bowden, M.G., Davis, S., Gurusiddappa, S., Moore, D., et al., A

dock, lock, and latch structural model for a staphylococcal adhesin binding to fibrinogen. Cell, 2003. **115**(2): p. 217-228.

[88] Bowden, M.G., Heuck, A.P., Ponnuraj, K., Kolosova, E., Choe, D., et al., Evidence for the dock, lock, and latch ligand binding mechanism of the staphylococcal microbial surface component recognizing adhesive matrix molecules (MSCRAMM) SdrG. J Biol Chem, 2008. **283**(1): p. 638-647.

[89] Mulcahy, M.E., Geoghegan, J.A., Monk, I.R., O'Keeffe, K.M., Walsh, E.J., et al., Nasal colonisation by *Staphylococcus aureus* depends upon clumping factor B binding to the squamous epithelial cell envelope protein loricrin. PLoS Pathog, 2012. **8**(12): p. e1003092. doi: 10.1371/journal.ppat.1003092

[90] Entenza, J.M., Foster, T.J., Ni Eidhin, D., Vaudaux, P., Francioli, P., et al., Contribution of clumping factor B to pathogenesis of experimental endocarditis due to *Staphylococcus aureus.* Infect Immun, 2000. **68**(9): p. 5443-5446.

[91] Moreillon, P., Entenza, J.M., Francioli, P., McDevitt, D., Foster, T.J., et al., Role of *Staphylococcus aureus* coagulase and clumping factor in pathogenesis of experimental endocarditis. Infect Immun, 1995. **63**(12): p. 4738-4743.

[92] Sinha, B., Francois, P.P., Nusse, O., Foti, M., Hartford, O.M., et al., Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin alpha5beta1. Cell Microbiol, 1999. **1**(2): p. 101-117.

[93] Flick, M.J., Du, X., Prasad, J.M., Raghu, H., Palumbo, J.S., et al., Genetic elimination of the binding motif on fibrinogen for the *S. aureus* virulence factor ClfA improves host survival in septicemia. Blood, 2013. **121**(10): p. 1783-1794.

[94] Cheng, A.G., Kim, H.K., Burts, M.L., Krausz, T., Schneewind, O., et al., Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. The FASEB Journal, 2009. **23**(10): p. 3393-3404.

[95] Hair, P.S., Echague, C.G., Sholl, A.M., Watkins, J.A., Geoghegan, J.A., et al., Clumping factor A interaction with complement factor I increases C3b cleavage on the bacterial surface of *Staphylococcus aureus* and decreases complement-mediated phagocytosis. Infect and Immun, 2010. **78**(4): p. 1717-1727.

[96] Josefsson, E., Hartford, O., O'Brien, L., Patti, J.M., and Foster, T., Protection against experimental *Staphylococcus aureus* arthritis by vaccination with clumping factor A, a novel virulence determinant. J Infect Dis, 2001. **184**(12): p. 1572-1580.

[97] O'Brien, L.M., Walsh, E.J., Massey, R.C., Peacock, S.J., and Foster, T.J., *Staphylococcus aureus* clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. Cell Microbiol, 2002. **4**(11): p. 759-770.

[98] Walsh, E.J., O'Brien, L.M., Liang, X., Hook, M., and Foster, T.J., Clumping Factor B, a fibrinogen-binding MSCRAMM (Microbial Surface Components Recognizing Adhesive Matrix Molecules) adhesin of *Staphylococcus aureus*, also binds to the tail region of Type I cytokeratin 10. J Biol Chem, 2004. **279**(49): p. 50691-50699.

[99] Schwarz-Linek, U., Werner, J.M., Pickford, A.R., Gurusiddappa, S., Kim, J.H., et al., Pathogenic bacteria attach to human fibronectin through a tandem beta-

zipper. Nature, 2003. **423**(6936): p. 177-181.

[100] Roche, F.M., Downer, R., Keane, F., Speziale, P., Park, P.W., et al., The N-terminal A domain of fibronectin-binding proteins A and B promotes adhesion of *Staphylococcus aureus* to elastin. J Biol Chem, 2004. **279**(37): p. 38433-38440.

[101] Vergara-Irigaray, M., Valle, J., Merino, N., Latasa, C., Garcia, B., et al., Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. Infect Immun, 2009. **77**(9): p. 3978-3991.

[102] Patti, J.M., Bremell, T., Krajewska-Pietrasik, D., Abdelnour, A., Tarkowski, A., et al., The *Staphylococcus aureus* collagen adhesin is a virulence determinant in experimental septic arthritis. Infect Immun, 1994. **62**(1): p. 152-161.

[103] Kang, M., Ko, Y. P., Liang, X., Ross, C.L., Liu, Q., et al., Collagen-binding Microbial Surface Components Recognizing Adhesive Matrix Molecule (MSCRAMM) of gram-positive bacteria inhibit complement activation via the classical pathway. J Biol Chem, 2013. **288**(28): p. 20520-20531.

[104] Barbu, E.M., Ganesh, V.K., Gurusiddappa, S., Mackenzie, R.C., Foster, T.J., et al., Beta-Neurexin is a ligand for the *Staphylococcus aureus* MSCRAMM SdrC. PLoS Pathog, 2010. **6**(1): p. e1000726. doi: 10.1371/journal.ppat.1000726

[105] Barbu, E.M., Mackenzie, C., Foster, T.J., and Hook, M., SdrC induces staphylococcal biofilm formation through a homophilic interaction. Mol Microbiol, 2014. **94**(1): p. 172-185.

[106] Corrigan, R.M., Miajlovic, H., and Foster, T.J., Surface proteins that promote adherence of *Staphylococcus aureus* to human desquamated nasal epithelial cells.

BMC Microbiol, 2009. **9**(1): p. 22-32. doi: 10.1186/1471-2180-9-22

[107] Zhang, Y., Wu, M., Hang, T., Wang, C., Yang, Y., et al., *Staphylococcus aureus* SdrE captures complement factor H's C-terminus via a novel 'close, dock, lock and latch' mechanism for complement evasion. Biochem J, 2017. **474**(10): p. 1619-1631.

[108] Vazquez, V., Liang, X., Horndahl, J.K., Ganesh, V.K., Smeds, E., et al., Fibrinogen is a ligand for the *Staphylococcus aureus* microbial surface components recognizing adhesive matrix molecules (MSCRAMM) bone sialoprotein-binding protein (Bbp). J Biol Chem, 2011. **286**(34): p. 29797-29805.

[109] Conlan, S., Mijares, L.A., Program, N.C.S., Becker, J., Blakesley, R.W., et al., *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. Genome Biol, 2012. **13**(7): p. R64. doi: 10.1186/gb-2012-13-7-r64

[110] Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A., Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol, 2005. **3**(9): p. 722-732.

[111] Jain, R., Rivera, M.C., Moore, J.E., and Lake, J.A., Horizontal gene transfer in microbial genome evolution. Theor Popul Biol, 2002. **61**(4): p. 489-495.

[112] Malachowa, N. and DeLeo, F.R., Mobile genetic elements of *Staphylococcus aureus.* Cell Mol Life Sci, 2010. **67**(18): p. 3057-3071.

[113] Weigel, L.M., Clewell, D.B., Gill, S.R., Clark, N.C., McDougal, L.K., et al., Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus*

aureus. Science, 2003. **302**(5650): p. 1569-1571.

[114] Schwarz, S., Gregory, P.D., Werckenthin, C., Curnock, S., and Dyke, K.G., A novel plasmid from *Staphylococcus epidermidis* specifying resistance to kanamycin, neomycin and tetracycline. J Med Microbiol, 1996. **45**(1): p. 57-63.

[115] Lampson, B.C. and Parisi, J.T., Nucleotide sequence of the constitutive macrolide-lincosamide-streptogramin B resistance plasmid pNE131 from *Staphylococcus epidermidis* and homologies with *Staphylococcus aureus* plasmids pE194 and pSN2. J Bacteriol, 1986. **167**(3): p. 888-892.

[116] Hatfull, G.F. and Hendrix, R.W., Bacteriophages and their Genomes. Curr Opin Virol, 2011. **1**(4): p. 298-303.

[117] Salmond, G.P. and Fineran, P.C., A century of the phage: past, present and future. Nat Rev Microbiol, 2015. **13**(12): p. 777-786.

[118] Holden, M.T., Lindsay, J.A., Corton, C., Quail, M.A., Cockfield, J.D., et al., 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, 2010. **192**(3): p. 888-892.

[119] Li, M., Du, X., Villaruz, A.E., Diep, B.A., Wang, D., et al., MRSA epidemic linked to a quickly spreading colonization and virulence determinant. Nat Med, 2012. **18**(5): p. 816-819.

[120] Rice, L.B., Tn916 Family conjugative transposons and dissemination of antimicrobial resistance derminants. Antimicrob Agents Chemother, 1998. **42**(8): p. 1871-1877.

[121] Poyart-Salmeron, C., Trieu-Cuot, P., Carlier, C., and Courvalin, P., Nucleotide sequences specific for Tn1545-like conjugative transposons in pneumococci and staphylococci resistant to tetracycline. Antimicrob Agents Chemother, 1991. **35**(8): p. 1657-1660.

[122] Roberts, A.P. and Mullany, P., A modular master on the move: the Tn916 family of mobile genetic elements. Trends Microbiol, 2009. **17**(6): p. 251-258.

[123] Rowland, S.J. and Dyke, K.G., Tn552, a novel transposable element from *Staphylococcus aureus.* Mol Microbiol, 1990. **4**(6): p. 961-975.

[124] Sidhu, M.S., Heir, E., Leegaard, T., Wiger, K., and Holck, A., Frequency of Disinfectant Resistance Genes and Genetic Linkage with β-Lactamase Transposon Tn552 among Clinical Staphylococci. Antimicrob Agents Chemother, 2002. **46**(9): p. 2797-2803.

[125] Juhas, M., Van der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W., et al., Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev, 2009. **33**(2): p. 376-393.

[126] Langille, M.G.I., Hsiao, W.W.L., and Brinkman, F.S.L., Detecting genomic islands using bioinformatics approaches. Nat Rev Micro, 2010. **8**(5): p. 373-382.

[127] Gill, S.R., Fouts, D.E., Archer, G.L., Mongodin, E.F., DeBoy, R.T., et al., 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, 2005. **187**(7): p. 2426-2438.

[128] Diep, B.A., Gill, S.R., Chang, R.F., Phan, T.H., Chen, J.H., et al., Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus.* Lancet, 2006. **367**(9512): p. 731-739.

[129] Shore, A.C., Rossney, A.S., Brennan, O.M., Kinnevey, P.M., Humphreys, H., et al., 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, 2011. **55**(5): p. 1896-1905.

[130] Monnet, V., Bacterial oligopeptide-binding proteins. Cel Mol Life Sci, 2003. **60**(10): p. 2100-2114.

[131] Cunin, R., Glansdorff, N., Piérard, A., and Stalon, V., Biosynthesis and metabolism of arginine in bacteria. Microbiol Rev, 1986. **50**(3): p. 314-352.

[132] Urushibara, N., Kawaguchiya, M., and Kobayashi, N., Two novel arginine catabolic mobile elements and staphylococcal chromosome cassette *mec* composite islands in community-acquired methicillin-resistant *Staphylococcus aureus* genotypes ST5-MRSA-V and ST5-MRSA-II. J Antimicrob Chemother, 2012. **67**(8): p. 1828-1834.

[133] Montgomery, C.P., Boyle-Vavra, S., and Daum, R.S., The arginine catabolic mobile element is not associated with enhanced virulence in experimental invasive disease caused by the community-associated methicillin-resistant *Staphylococcus aureus* USA300 genetic background. Infect Immun, 2009. **77**(7):

p. 2650-2656.

[134] Diep, B.A., Stone, G.G., Basuino, L., Graber, C.J., Miller, A., et al., 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, 2008. **197**(11): p. 1523-1530.

[135] Onishi, M., Urushibara, N., Kawaguchiya, M., Ghosh, S., Shinagawa, M., et al., 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, 2013. **20**: p. 381-388.

[136] Miragaia, M., de Lencastre, H., Perdreau-Remington, F., Chambers, H.F., Higashi, J., et al., Genetic diversity of arginine catabolic mobile element in *Staphylococcus epidermidis.* PLoS One, 2009. **4**(11):e7722. doi: 10.1371/journal.pone.0007722.

[137] Bartels, M.D., Hansen, L.H., Boye, K., Sørensen, S.J., and Westh, H., An unexpected location of the arginine catabolic mobile element (ACME) in a USA300-related MRSA strain. PLoS One, 2011. **6**(1): p. e16193. doi: 10.1371/journal.pone.0016193

[138] Hiramatsu, K., Ito, T., Tsubakishita, S., Sasaki, T., Takeuchi, F., et al., Genomic basis for methicillin resistance in *Staphylococcus aureus.* Infect Chemother, 2013. **45**(2): p. 117-136.

[139] Scheffers, D.J. and Pinho, M.G., Bacterial cell wall synthesis: new insights from

localization studies. Microbiol Mol Biol Rev, 2005. **69**(4): p. 585-607.

[140] Pinho, M.G., Kjos, M., and Veening, J.W., How to get (a)round: mechanisms controlling growth and division of coccoid bacteria. Nat Rev Microbiol, 2013. **11**(9): p. 601-614.

[141] Tipper, D.J. and Strominger, J.L., Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. Proc Natl Acad Sci U S A, 1965. **54**(4): p. 1133-1141.

[142] Peacock, S.J. and Paterson, G.K., Mechanisms of Methicillin Resistance in *Staphylococcus aureus.* Annu Rev Biochem, 2015. **84**: p. 577-601.

[143] Fuda, C., Suvorov, M., Vakulenko, S.B., and Mobashery, S., The basis for resistance to beta-lactam antibiotics by penicillin-binding protein 2a of methicillin-resistant *Staphylococcus aureus.* J Biol Chem, 2004. **279**(39): p. 40802-40806.

[144] Lim, D. and Strynadka, N.C., Structural basis for the beta lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus.* Nat Struct Biol, 2002. **9**(11): p. 870-876.

[145] Misiura, A., Pigli, Y.Z., Boyle-Vavra, S., Daum, R.S., Boocock, M.R., et al., Roles of two large serine recombinases in mobilizing the methicillin-resistance cassette *SCCmec.* Mol Microbiol, 2013. **88**(6): p. 1218-1229.

[146] Fagerlund, A., Granum, P.E., and Havarstein, L.S., *Staphylococcus aureus* competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. Mol Microbiol, 2014. **94**(3): p. 557-579.

[147] Furuya, E.Y. and Lowy, F.D., Antimicrobial-resistant bacteria in the community setting. Nat Rev Microbiol, 2006. **4**(1): p. 36-45.

[148] Lindsay, J.A., *Staphylococcus aureus* genomics and the impact of horizontal gene transfer. Int J Med Microbiol, 2014. **304**(2): p. 103-109.

[149] Scharn, C.R., Tenover, F.C., and Goering, R.V., Transduction of staphylococcal cassette chromosome mec elements between strains of *Staphylococcus aureus.* Antimicrob Agents Chemother, 2013. **57**(11): p. 5233-5238.

[150] Otto, M., 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, 2013. **35**(1): p. 4-11.

[151] Wu, S., Piscitelli, C., de Lencastre, H., and Tomasz, A., 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, 1996. **2**(4): p. 435-441.

[152] Barbier, F., Ruppe, E., Hernandez, D., Lebeaux, D., Francois, P., et al., Methicillin-resistant coagulase-negative staphylococci in the community: high homology of SCC*mec* IVa between *Staphylococcus epidermidis* and major clones of methicillin-resistant *Staphylococcus aureus.* J Infect Dis, 2010. **202**(2): p. 270-281.

[153] Archer, G.L., Thanassi, J.A., Niemeyer, D.M., and Pucci, M.J., Characterization of IS1272, an insertion sequence-like element from *Staphylococcus*

haemolyticus. Antimicrob Agents Chemother, 1996. **40**(4): p. 924-929.

[154] Wisplinghoff, H., Rosato, A.E., Enright, M.C., Noto, M., Craig, W., et al., Related clones containing SCC*mec* Type IV predominate among clinically significant *Staphylococcus epidermidis* isolates. Antimicrob Agents Chemother, 2003. **47**(11): p. 3574-3579.

[155] Wallmark, G., Arremark, I., and Telander, B., *Staphylococcus saprophyticus*: a frequent cause of acute urinary tract infection among female outpatients. J Infect Dis, 1978. **138**(6): p. 791-797.

[156] Van Der Zwet, W.C., Debets-Ossenkopp, Y.J., Reinders, E., Kapi, M., Savelkoul, P.H., et al., Nosocomial spread of a *Staphylococcus capitis* strain with heteroresistance to vancomycin in a neonatal intensive care unit. J Clin Microbiol, 2002. **40**(7): p. 2520-2525.

[157] Gras-Le Guen, C., Fournier, S., Andre-Richet, B., Caillon, J., Chamoux, C., et al., Almond oil implicated in a *Staphylococcus capitis* outbreak in a neonatal intensive care unit. J Perinatol, 2007. **27**(11): p. 713-717.

[158] Petti, C.A., Simmon, K.E., Miro, J.M., Hoen, B., Marco, F., et al., Genotypic diversity of coagulase-negative staphylococci causing endocarditis: a global perspective. J Clin Microbiol, 2008. **46**(5): p. 1780-1784.

[159] Foster, T.J., Molecular genetic analysis of staphylococcal virulence. Method Microbiol, 1998. **27**: p. 434- 454.

[160] Schaeffer, C.R., Woods, K.M., Longo, G.M., Kiedrowski, M.R., Paharik, A.E., et al., Accumulation-associated protein enhances *Staphylococcus epidermidis*

biofilm formation under dynamic conditions and is required for infection in a rat catheter model. Infect Immun, 2015. **83**(1): p. 214-226.

[161] Christner, M., Franke, G., Schommer, N., Wendt, U., Wegert, K., et al., The giant extracellular matrix binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment to fibronectin. Mol Microbiol, 2010. **75**(1): 187-207. doi: 10.1111/j.1365-2958.2009.06981.x

[162] Shahrooei, M., Hira, V., Stijlemans, B., Merckx, R., Hermans, P.W., et al., Inhibition of *Staphylococcus epidermidis* biofilm formation by rabbit polyclonal antibodies against the SesC protein. Infect Immun, 2009. **77**(9): p. 3670-3678.

[163] Ramboarina, S., Garnett, J.A., Zhou, M., Li, Y., Peng, Z., et al., Structural insights into serine-rich fimbriae from Gram-positive bacteria. J Biol Chem, 2010. **285**(42): p. 32446-32457.

[164] Pyburn, T.M., Bensing, B.A., Xiong, Y.Q., Melancon, B.J., Tomasiak, T.M., et al., A structural model for binding of the serine-rich repeat adhesin GspB to host carbohydrate receptors. PLoS Pathog, 2011. **7**(7): p. e1002112. doi: 10.1371/journal.ppat.1002112

[165] Siboo, I.R., Chambers, H.F., and Sullam, P.M., Role of SraP, a serine-rich surface protein of *Staphylococcus aureus*, in binding to human platelets. Infect Immun, 2005. **73**(4): p. 2273-2280.

[166] Zhou, M. and Wu, H., Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. Microbiology, 2009. **155**(Pt 2): p. 317-327.

[167] Schneider, T.D. and Stephens, R.M., Sequence logos: a new way to display

consensus sequences. Nucleic Acids Res, 1990. **18**(20): p. 6097-6100.

[168] Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E., WebLogo: a sequence logo generator. Genome Res, 2004. **14**(6): p. 1188-1190.

[169] Gordon, R.J., Weinberg, A.D., Pagani, F.D., Slaughter, M.S., Pappas, P.S., et al., Prospective, multicenter study of ventricular assist device infections. Circulation, 2013. **127**(6): p. 691-702.

[170] Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J., The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc, 2015. **10**(6): p. 845-858.

[171] Josefsson, E., O'Connell, D., Foster, T.J., Durussel, I., and Cox, J.A., The binding of calcium to the B-repeat segment of SdrD, a cell surface protein of *Staphylococcus aureus.* J Biol Chem, 1998. **273**(47): p. 31145-31152.

[172] Wang, X., Ge, J., Liu, B., Hu, Y., and Yang, M., Structures of SdrD from *Staphylococcus aureus* reveal the molecular mechanism of how the cell surface receptors recognize their ligands. Protein Cell, 2013. **4**(4): p. 277-285.

[173] Gruszka, D.T., Wojdyla, J.A., Bingham, R.J., Turkenburg, J.P., Manfield, I.W., et al., Staphylococcal biofilm-forming protein has a contiguous rod-like structure. Proc Natl Acad Sci U S A, 2012. **109**(17): p. E1011-1018.

[174] Conrady, D.G., Wilson, J.J., and Herr, A.B., Structural basis for Zn2+-dependent intercellular adhesion in staphylococcal biofilms. Proc Natl Acad Sci U S A, 2013. **110**(3): p. E202-211.

[175] Sakinc, T., Kleine, B., and Gatermann, S.G., SdrI, a serine-aspartate repeat

protein identified in *Staphylococcus saprophyticus* strain 7108, is a collagen-binding protein. Infect Immun, 2006. **74**(8): p. 4615-4623.

[176] Hazenbos, W.L., Kajihara, K.K., Vandlen, R., Morisaki, J.H., Lehar, S.M., et al., Novel staphylococcal glycosyltransferases SdgA and SdgB mediate immunogenicity and protection of virulence-associated cell wall proteins. PLoS Pathog, 2013. **9**(10): p. e1003653. doi: 10.1371/journal.ppat.1003653

[177] Ythier, M., Resch, G., Waridel, P., Panchaud, A., Gfeller, A., et al., Proteomic and transcriptomic profiling of *Staphylococcus aureus* surface LPXTG-proteins: correlation with agr genotypes and adherence phenotypes. Mol Cell Proteomics, 2012. **11**(11): p. 1123-1139.

[178] Sellman, B.R., Timofeyeva, Y., Nanra, J., Scott, A., Fulginiti, J.P., et al., Expression of *Staphylococcus epidermidis* SdrG increases following exposure to an in vivo environment. Infect Immun, 2008. **76**(7): p. 2950-2957.

[179] Di Poto, A., Papi, M., Trivedi, S., Maiorana, A., Gavazzo, P., et al., In vitro effect of temperature on the conformational structure and collagen binding of SdrF, a *Staphylococcus epidermidis* adhesin. Appl Microbiol Biotechnol, 2015. **99**(13): p. 5593-5603.

[180] Conlon, B.P., Geoghegan, J.A., Waters, E.M., McCarthy, H., Rowe, S.E., et al., Role for the A domain of unprocessed accumulation-associated protein (Aap) in the attachment phase of the *Staphylococcus epidermidis* biofilm phenotype. J Bacteriol, 2014. **196**(24): p. 4268-4275.

[181] Kozitskaya, S., Olson, M.E., Fey, P.D., Witte, W., Ohlsen, K., et al., Clonal

analysis of *Staphylococcus epidermidis* isolates carrying or lacking biofilm-mediating genes by multilocus sequence typing. J Clin Microbiol, 2005. **43**(9): p. 4751-4757.

[182] Schoenfelder, S.M., Lange, C., Eckart, M., Hennig, S., Kozytska, S., et al., Success through diversity - how *Staphylococcus epidermidis* establishes as a nosocomial pathogen. Int J Med Microbiol, 2010. **300**(6): p. 380-386.

[183] Miragaia, M., Thomas, J.C., Couto, I., Enright, M.C., and de Lencastre, H., Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing data. J Bacteriol, 2007. **189**(6): p. 2540-2552.

[184] Zong, Z., Peng, C., and Lu, X., Diversity of SCC*mec* elements in methicillin-resistant coagulase-negative staphylococci clinical isolates. PLoS One, 2011. **6**(5): p. e20191. doi: 10.1371/journal.pone.0020191

[185] Garza-Gonzalez, E., Lopez, D., Pezina, C., Muruet, W., Bocanegra-Garcia, V., et al., Diversity of staphylococcal cassette chromosome mec structures in coagulase-negative staphylococci and relationship to drug resistance. J Med Microbiol, 2010. **59**(Pt 3): p. 323-329.

[186] Mombach P.M.A.B., Reiter, K.C., Paiva, R.M., and Barth, A.L., Distribution of staphylococcal cassette chromosome *mec* (SCC*mec*) types I, II, III and IV in coagulase-negative staphylococci from patients attending a tertiary hospital in southern Brazil. J Med Microbiol, 2007. **56**(Pt 10): p. 1328-1333.

[187] McManus, B.A., Coleman, D.C., Deasy, E.C., Brennan, G.I., O'Connell, B., et al., 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, 2015. **10**(9): p. e0138079. doi: 10.1371/journal.pone.0138079

[188] Ito, T., Hiramatsu, K., Oliveira, D.C., deLencastre, H., Zhang, K., et al., Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. Antimicrob Agents Chemother, 2009. **53**(12): 4961-4967.

[189] Mkrtchyan, H.V., Xu, Z., and Cutler, R.R., Diversity of SCC*mec* elements in staphylococci isolated from public washrooms. BMC Microbiol, 2015. **15**: p. 120-125. doi: 10.1186/s12866-015-0451-3

[190] Ibrahem, S., Salmenlinna, S., Lyytikainen, O., Vaara, M., and Vuopio-Varkila, J., Molecular characterization of methicillin-resistant *Staphylococcus epidermidis* strains from bacteraemic patients. Clin Microbiol Infect, 2008. **14**(11): p. 1020-1027.

[191] Najar-Peerayeh, S., Moghaddas, A.J., Bakhshi, B., and Ghasemian, A., Diversity of the SCC*mec* types among *Staphylococcus epidermidis* clinical isolates from intensive care unit patients. Asian Pac J Trop Dis, 2016. **6**(2): p. 133-135.

[192] Monecke, S., Coombs, G.W., Pearson, J., Hotzel, H., Slickers, P., et al., A clonal complex 12 methicillin-resistant *Staphylococcus aureus* strain, West Australian MRSA-59, harbors a novel pseudo-SCC*mec* element. Antimicrob Agents Chemother, 2015. **59**(11): p. 7142-7144.

[193] Perreten, V., Chanchaithong, P., Prapasarakul, N., Rossano, A., Blum, S.E., et al., Novel pseudo-staphylococcal cassette chromosome *mec* element (psiSCC*mec*57395) in methicillin-resistant *Staphylococcus pseudintermedius* CC45. Antimicrob Agents Chemother, 2013. **57**(11): p. 5509-5515.

[194] Okuma, K., Iwakawa, K., Turnidge, J.D., Grubb, W.B., Bell, J.M., et al., Dissemination of new methicillin-resistant *Staphylococcus aureus* clones in the community. J Clin Microbiol, 2002. **40**(11): p. 4289-4294.

[195] Heusser, R., Ender, M., Berger-Bachi, B., and McCallum, N., Mosaic staphylococcal cassette chromosome *mec* containing two recombinase loci and a new *mec* complex, B2. Antimicrob Agents Chemother, 2007. **51**(1): p. 390-393.

[196] Luong, T.T., Ouyang, S., Bush, K., and Lee, C.Y., Type 1 capsule genes of *Staphylococcus aureus* are carried in a staphylococcal cassette chromosome genetic element. J Bacteriol, 2002. **184**(13): p. 3623-3629.

[197] Chongtrakool, P., Ito, T., Ma, X.X., Kondo, Y., Trakulsomboon, S., et al., 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, 2006. **50**(3): p. 1001-1012.

[198] International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements, Classification of staphylococcal cassette chromosome mec (SCC*mec*): guidelines for reporting novel SCC*mec* elements. Antimicrob Agents Chemother, 2009. **53**(12): p. 4961-4967.

[199] Zhang, X., Wu, M., Zhuo, W., Gu, J., Zhang, S., et al., Crystal structures of Bbp from *Staphylococcus aureus* reveal the ligand binding mechanism with Fibrinogen alpha. Protein Cell, 2015. **6**(10): p. 757-766.

[200] Stemberk, V., Jones, R.P., Moroz, O., Atkin, K.E., Edwards, A.M., et al., Evidence for steric regulation of fibrinogen binding to *Staphylococcus aureus* fibronectin-binding protein A (FnBPA). J Biol Chem, 2014. **289**(18): p. 12842-12851.

[201] Arora, S., Uhlemann, A.C., Lowy, F.D., and Hook, M., A novel MSCRAMM subfamily in coagulase negative staphylococcal Species. Front Microbiol, 2016. **7**: p. 540-549. doi: 10.3389/fmicb.2016.00540

[202] Savolainen, K., Paulin, L., Westerlund-Wikstrom, B., Foster, T.J., Korhonen, T.K., et al., Expression of *pls*, a gene closely associated with the *mecA* gene of methicillin-resistant *Staphylococcus aureus, prevents bacterial adhesion in vitro.* Infect Immun, 2001. **69**(5): p. 3013-3020.

[203] Ito, T., Katayama, Y., Asada, K., Mori, N., Tsutsumimoto, K., et al., Structural comparison of three types of staphylococcal cassette chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus.* Antimicrob Agents Chemother, 2001. **45**(5): p. 1323-1336.

[204] Bleiziffer, I., Eikmeier, J., Pohlentz, G., McAulay, K., Xia, G., et al., The plasmin-sensitive Ppotein Pls in methicillin-resistant *Staphylococcus aureus* (MRSA) is a glycoprotein. PLoS Pathog, 2017. **13**(1): p. e1006110.

[205] Huesca, M., Peralta, R., Sauder, D.N., Simor, A.E., and McGavin, M.J.,

Adhesion and virulence properties of epidemic Canadian methicillin-resistant *Staphylococcus aureus* strain 1: identification of novel adhesion functions associated with plasmin-sensitive surface protein. J Infect Dis, 2002. **185**(9): p. 1285-1296.

[206] Josefsson, E., Juuti, K., Bokarewa, M., and Kuusela, P., The surface protein Pls of methicillin-resistant *Staphylococcus aureus* is a virulence factor in septic arthritis. Infect Immun, 2005. **73**(5): p. 2812-2817.

[207] Jolley, K.A. and Maiden, M.C., BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics, 2010. **11**: p. 595-606. doi: 10.1186/1471-2105-11-595

[208] Nascimento, M., Sousa, A., Ramirez, M., Francisco, A.P., Carrico, J.A., et al., PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. Bioinformatics, 2017. **33**(1): p. 128-129.

[209] Bolger, A.M., Lohse, M., and Usadel, B., Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 2014. **30**(15): p. 2114-2120.

[210] Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., et al., SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol, 2012. **19**(5): p. 455-477.

[211] Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., et al., Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods, 2013. **10**(6): p. 563-569.

[212] Langmead, B. and Salzberg, S.L., Fast gapped-read alignment with Bowtie 2. Nat Methods, 2012. **9**(4): p. 357-359.

[213] Loughman, A., Sweeney, T., Keane, F.M., Pietrocola, G., Speziale, P., et al., Sequence diversity in the A domain of *Staphylococcus aureus* fibronectin-binding protein A. BMC Microbiol, 2008. **8**: p. 74-88. doi: 10.1186/1471-2180-8-74

[214] Wilson, L.K., Coombs, G.W., Christiansen, K., Grubb, W.B., and O'Brien, F.G., Characterization of a novel staphylococcal cassette chromosome composite island from community-associated MRSA isolated in aged care facilities in Western Australia. J Antimicrob Chemother, 2016. **71**(12): p. 3372-3375.

[215] Du, X., Zhu, Y., Song, Y., Li, T., Luo, T., et al., Molecular analysis of *Staphylococcus epidermidis* strains isolated from community and hospital environments in China. PLoS One, 2013. **8**(5): p. e62742.

[216] Widerstrom, M., McCullough, C.A., Coombs, G.W., Monsen, T., and Christiansen, K.J., A multidrug-resistant *Staphylococcus epidermidis* clone (ST2) is an ongoing cause of hospital-acquired infection in a Western Australian hospital. J Clin Microbiol, 2012. **50**(6): p. 2147-2151.

[217] Deplano, A., Vandendriessche, S., Nonhoff, C., Dodemont, M., Roisin, S., et al., National surveillance of *Staphylococcus epidermidis* recovered from bloodstream infections in Belgian hospitals. J Antimicrob Chemother, 2016. **71**(7): p. 1815-1819.

[218] Cremniter, J., Sivadon-Tardy, V., Caulliez, C., Bauer, T., Porcher, R., et al.,

Genetic analysis of glycopeptide-resistant *Staphylococcus epidermidis* strains from bone and joint infections. J Clin Microbiol, 2013. **51**(3): p. 1014-1019.

[219] Brenciani, A., Morroni, G., Pollini, S., Tiberi, E., Mingoia, M., et al., Characterization of novel conjugative multiresistance plasmids carrying cfr from linezolid-resistant *Staphylococcus epidermidis* clinical isolates from Italy. J Antimicrob Chemother, 2016. **71**(2): p. 307-313.

[220] Burke, F.M., McCormack, N., Rindi, S., Speziale, P., and Foster, T.J., Fibronectin-binding protein B variation in *Staphylococcus aureus.* BMC Microbiol, 2010. **10**: p. 160-175. doi: 10.1186/1471-2180-10-160

[221] Shore, A.C., Rossney, A.S., O'Connell, B., Herra, C.M., Sullivan, D.J., et al., Detection of staphylococcal cassette chromosome *mec*-associated DNA segments in multiresistant methicillin-susceptible *Staphylococcus aureus* (MSSA) and identification of *Staphylococcus epidermidis* ccrAB4 in both methicillin-resistant *S. aureus* and MSSA. Antimicrob Agents Chemother, 2008. **52**(12): p. 4407-4419.

[222] McCarthy, H., Rudkin, J.K., Black, N.S., Gallagher, L., O'Neill, E., et al., Methicillin resistance and the biofilm phenotype in S*taphylococcus aureus.* Front Cell Infect Microbiol, 2015. **5**: p. 1-7. doi: 10.3389/fcimb.2015.00001