Comparative genomics of environmental multidrug resistant staphylococci

Rory Joseph Cave

Ph.D.

September 2019

Abstract

Multidrug resistant (MDR) staphylococci are public health concern internationally. The treatment of these bacteria have become increasingly difficult due to their resistance to multiple antibiotics. In this thesis, multidrug resistant staphylococci have been recovered from high-frequency touched surfaces in public areas in the community and hospitals in East and West London. In total, 600 isolates collected of which 281 were MDR. In addition, 49 (8.17%) were *mecA* gene positive (*mecA*$^+$). The most common species identified as multidrug resistant were *S. epidermidis*, *S. haemolyticus* and *S. hominis*, whereas penicillin, fusidic acid and erythromycin were the most frequent antibiotics the isolates were resistant to. Whole genome sequencing (WGS) analysis for *mecA*$^+$ isolates revealed that among the most frequent antibiotic resistance genes were *blaZ*, *qacA/B* and *dfrC*. Moreover, the *mecA*$^+$ isolates had a diverse range of *SCCmec* types many of which were untypable due to carrying a novel combination of *ccr* genes or multiple *ccr* complexes. *mecA*$^+$ *S. epidermidis*, *S. haemolyticus* and *S. hominis* isolates that have been whole genome sequenced were used in the "One Health" comparative genomics approach to compare them with isolates obtained from the ENA database that were recovered from clinical samples, healthy human body sites, livestock, companion animals and other environments. The *mecA*$^+$ *S. epidermidis* and *S. haemolyticus* isolates in this study were genetically related and shared similar accessory gene profiles with ENA isolates that have been recovered from clinical samples. In addition, all three species *mecA*$^+$ isolates recovered from public settings were genetically related to ENA isolates recovered from different source including healthy humans, livestock, and companion animals, plants and other environmental sources. In conclusion, the high-frequency touched surfaces in public settings are reservoirs for staphylococci belonging to different lineages that are multidrug resistant and therefore pose a potential public health risk.

Table of contents

List of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 6

Appendix

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Abbreviations

AMR……..……...Antimicrobial resistance

CA-MRSA…….…Community-associated Methicillin resistance *Staphylococcus aureus*

CoNS…….…….……Coagulase negative staphylococci

CoPS……...……Coagulase positive staphylococci

ELC…….……...……. East London community

ELH…….…..……East London Hospital

ENA……...……..……European Nucleotide Archive

GC……...……...……Guanine-Cytosine

HA-MRSA……..…. Hospital-associated Methicillin resistance *Staphylococcus aureus*

HGT……...Horizontal gene transfer

LA-MRSA….……Livestock-associated Methicillin resistance *Staphylococcus aureus*

MALDI-TOF……. Matrix-Assisted Laser Desorption/Ionization-Time of Flight

MLST……...……...…Multilocus sequence type

MRSA……..……..…..Methicillin resistance *Staphylococcus aureus*

MDR……...……Multidrug resistance

NCBI……...…….National Centre for Biotechnology and technology

PCR……...…… Polymerase chain reaction

PFGE……...……..Pulsed field Gel electrophoresis

ST……..Sequence type

t-SNE… …….……..t-Distributed Stochastic Neighbour Embedding

WGS…….……Whole genome sequencing

WLC……...…….……West London community

WLH……...……...……West London hospital

Acknowledgement

Throughout my PhD research project and thesis, I have received a great deal of support and assistance. I would like to thank my director of studies Dr Hermine Mkrtchyan for giving me this opportunity to conduct this research project, her guidance and time she spent reading through my thesis.

Many thanks to my other two supervisors, Dr Raju Misra, for his expertise in bioinformatics, and Dr David Bringloe, for his knowledge of how to navigate through the PhD process at the University of East London.

I would also like to thank Professor Haroun shah, and Professor Ajit Shah from Middlesex University for their collaboration and letting me use their MALDI-TOF instruments and given me technical support. Dr Jiazhen Chen and Shiyong Wang from Fudan University for whole genome sequencing my samples for me. My Dad; Dr Mark Cave for proofreading my thesis, his guidance on my statistical analyses and helping me troubleshooting any problems I had with the R programming language.

Special thanks to my fiancée Sha Zhou for motivating me to make self-improvement, supplying me with emotional support and distracting me from my thesis with hotpot.

And finally, I would like to thank the rest of my family and friends for the continued support, guidance and encouragement they have given me throughout my time as a PhD student. This work would not be completed without them.

Chapter 1: Introduction

Antimicrobial resistance (AMR) is a global problem (Ventola, 2015). Today many clinically important species have been identified to be multidrug resistant (MDR) (Alirol et al., 2017; Aloush et al., 2006; Gibson et al., 2010; Waters et al., 2011). This has been caused by the overuse of antibiotics (Michael et al., 2014). With the lack of development of new antibiotics, it is predicted there will be a post-antibiotic era where simple bacterial infection would lead to mortality as well as surgery no longer being plausible due to a risk of infections (Bragg et al., 2018). It has been predicted that by 2050, up to 10 million lives will be at risk (O'Neill, 2016). AMR bacteria are commonly found in clinical settings and animal agriculture due to the high amount of antibiotic used in these areas (Michael et al., 2014). From these areas, the AMR bacteria can disseminate into other niches (von Wintersdorff et al., 2016). One area in which AMR bacteria could spread and then further disseminate to other areas is in public settings. People can transmit AMR bacteria in these settings onto high-frequency hand touched surfaces (Bhatta et al., 2018; Roberts et al., 2013). The cycle can continue by further transmission from these surfaces onto people's hands and then transmitted to other surfaces, spreading the AMR bacteria (Bhatta et al., 2018). One genus of bacteria that has already shown to spread AMR into public settings are staphylococci (Conceição et al., 2013; Lutz et al., 2014; Seng et al., 2017b; Xu et al., 2015). Studies of MDR staphylococci from public settings are fragmentary with few reports mainly focusing on isolates recovered from buses, hotels, beaches and University campus (Conceição et al., 2013; Lutz et al., 2014, 2014; Seng et al., 2017a; Xu et al., 2015).

1. Literature review

1.1 Staphylococci characteristics

1.1.1 Morphology

Staphylococci are Gram-positive cocci with a diameter ranging from 0.5-1.0μm. They are seen to grow in "grape-like" cluster but occasionally have occurred singly, in pairs, tetrads and short chains (Foster, 1996). These organisms are non-motile with some members of this genus having a capsule. On agar plates, they usually form dwarf colonies (Borderon and Horodniceanu, 1978). These colonies can appear smooth, glossy, butyrous or wet. For the majority of the time, colonies are opaque, but some can be pigmented, which are creamy-white in colour or yellow/orange in colour (Becker et al., 2014).

1.1.2. Biochemical properties

Staphylococci are catalase-positive facultative anaerobes that can grow at temperatures ranging from 30ºC-37ºC. (Conceição et al., 2013; Lutz et al., 2014; Soge et al., 2009; Xu et al., 2015). They are also oxidase-negative (though a few species are oxidase-positive), urease-variable, ornithine-decarboxylase-negative and halotolerant (Scybert et al., 2003; Tsoi and Tse, 2011; Vernozy-Rozand et al., 2000; Zhou et al., 2019). Staphylococci can be split into two groups; coagulase-positive (CoPS) and coagulase-negative staphylococci (CoNS). CoPS produce coagulase an enzyme which causes blood clotting, whereas CoNS lack this enzyme (Sasaki et al., 2010). Their chromosomal guanine-cytosine (GC) content range from 30-40% (Suzuki et al., 2012).

1.1.3 Taxonomy

Staphylococci belong to the Firmicutes phylum; bacilli class; Bacillales order and the Staphylococcaceae family (Becker et al., 2014). *Staphylococcus* consists of 47 species and 23 subspecies (Table 1.1 and Table 1.2) (Becker et al., 2014). Eight of these species are CoPS and 38 are CoNS. *S. schleiferi* can be both CoPS (*S. schleiferi* subsp. *Coagulans*) and CoNS (*S. schleiferi* subsp. *Schleiferi*) (Becker et al., 2014). *Staphylococcus* can be

further grouped phylogenetically into 6 species group and 15 cluster groups (Becker et al., 2014).

Species group	Hycius-Intermidius			Epidermidis-Aureus				
Cluster group	Muscae	Hyicus	Intermedius	Aureus	Epidermidis	Warneri	Haemolyticus	Lugdunensis
Species	*S. muscae* *S. micorti* *S. rosti*	*S. hyicus*[2] *S. agenetis*[2] *S. chromogenes* *S. felis*	*S. intermedius*[1] *S. delphini*[1] *S. lutrae*[1] *S. pseudinter-medius*[1] *S. schleiferi* ssp. *schleiferi* ssp. *coagulans*[1]	*S. aureus* Ssp. aureus[1] ssp. anaerobius[1] *S. simiae*[1]	*S. epidermidis* *S. capitis* spp. *capitis* spp. *urealyticus* *S. caprae* *S. saccharolyticus*	*S. warneri* *S. pasteuri*	*S. haemolyticus* *S. devriesei* *S. hominis* spp. *hominis* spp. *novobio-septicus* *S. jettensis* *S. petrasii* Spp. *croceilytcicus* spp. *petasii*	*S. lugdunensis*

Table 1.1: Phylogenetic separation of *Staphylococcus* species part 1. 1= CoPS; 2= coagulase variable. Table from Becker et al., 2014.

Species group	Auricularis	Simulans	Saprophyticus				Sciuri
Cluster group	Auricularis	Simulans-Carnosus	Pettenkoferi-Massiliensis	Cohnii-Nepalensis	Saprophyticus	Arlettae-Kloosii	Sciuri
Species	*S. auricularis*	*S. simulans* *S. carnosus* spp. *carnosus* spp. *utilis* *S. condimenti* *S. piscifermentans*	*S. pettenkoferi* *S. massiliensis*	*S. cohnii* spp. *cohnii* spp. *urealyticus* *S. nepalensis*	*S. saprophyticus* spp. *saprophy-ticus* spp. *bovis* *S. equorum* spp. *equorum* spp. *linen* *S. gallinarum* *S. succinus* spp.. *succinus* *S. xylosus*	*S. arlettae* *S. kloosii*	*S. sciuri* spp. *sciuri* spp. *carnaticus* spp. *rodentium* *S. fleurettii* *S. lentus* *S. stepanovivii* *S. vitulinus*

Table 1.2: Phylogenetic separation of *Staphylococcus* species part 2. Table from Becker et al., 2014

1.1.4 Epidemiology and transmission

Staphylococci have initially been described as either being human-associated or animal-associated although studies have shown that members of these species can cross over to other species (Table 1.2) (Becker et al., 2014). They are found as part of the natural microbiota of the skin and mucous membrane of human and animals. Data shows that staphylococci can be found on axillar, the gluteal and inguinal regions, anterior nares, the umbilicus, the antecubital and popliteal spaces, the plantar foot regions, ocular surfaces and the conjunctiva (Costello et al., 2009; Graham et al., 2007; Grice et al., 2009; Wos-Oxley et al., 2010). As these organisms are found on the skin, they can be transmitted across from humans and animal and *vice versa* (Gómez-Sanz et al., 2019; Velasco et al., 2015). They can also be transmitted via an intermediate object that comes into contact with human or animal skin that contains these microorganisms (Conceição et al., 2013; Lutz et al., 2014; Seng et al., 2017b; Xu et al., 2015).

Species or subspecies	Site or source of detection		
	Environment and/or food	**Animals**	**Humans**
S. arlettae	Textile and tannery industrial effluents	Cattle, goats, pigs, poultry, sheep	–
S. aureus	Raw meat, public beaches, buses, housing and public and build-up areas	Cattle, goats, pigs, poultry, sheep, dogs	Skin, anterior nares
S. auricularis	–	–	External auditory canal (principle habitat), seldom on other skin regions
S. capitis subsp. *capitis*	–	Cats, dogs, horses	Predominantly on the scalp and arms, less frequently on other skin regions
S. capitis subsp. *urealyticus*	–	–	Predominantly on the skin (mostly from heads, primarily ears and foreheads)
S. caprae	–	Goats	Skin, anterior nares
*S. carnosus*subsp. *carnosus*	Fermented food (starter cultures, soy sauce mash)	Cattle	–
*S. carnosus*subsp. *utilis*	Fermented food (soy sauce mash, fermented fish)		–
S. chromogenes	–	Cattle, pigs, horses, goats, sheep	–

Species or subspecies	**Site or source of detection**		
S. cohnii subsp. *cohnii*	Hotel rooms, university campus	Dogs, goats, poultry	Skin
S. cohnii subsp. *urealyticus*	–	Apes, clams, monkeys, horses	Skin
S. condimenti	Fermented food and starter cultures	–	–
S. devriesei	–	Cattle	–
S. epidermidis	Fermented sausages, Hotel rooms, rice seeds, university campus	Cats, cattle, dogs, goats, gorillas, horses, pigs, sheep	Skin (preferentially axillae and the head; also arms and legs) and mucous membranes of the nasopharynx
S. equorum subsp. *equorum*	Fermented food (starter cultures)	Cattle, goats, horses, sheep	–
S. equorum subsp. *linens*	Smear-ripened cheese (starter culture)	–	–
S. felis	–	Cats, horses	–
S. fleurettii	Milk cheese	Goats, pigs, small mammals	–
S. gallinarum		Chickens, pheasants	–

Species or subspecies	Site or source of detection		
S. haemolyticus	Milk, fermented food, Hotel rooms, university campus	Cats, cattle, dogs, horses, goats, pigs, sheep	Skin (preferentially legs and arms)
S. hominis subsp. *hominis*	Goat milk, fermented food, Hotel rooms, university campus	Cats, dogs, goats, pigs, sheep	Skin (preferentially axillae, arms, legs, and pubic and inguinal regions)
S. hominis subsp. *novobiosepticus*	–	–	–
S. jettensis	–	–	–
S. kloosii	–	Goats	–
S. lentus	Soybean oil meal, meat, milk	Clams, goats, horses, mink, pigs, poultry, sheep	–
S. lugdunensis	–Hotel room	Cats, chinchillas, dogs, goats, guinea pigs	Skin (preferentially lower abdomen and extremities)
S. massiliensis	–	–	Skin
S. microti	–	Mice	–
S. muscae	–	Flies (trapped in cattle sheds)	–

Species or subspecies	Site or source of detection		
S. nepalensis	Environment (not specified)	Goats, pigs, squirrel monkeys, bats (guano), dry-cured ham	–
S. pasteuri	Fermented sausages	Pigs	–
S. petrasii subsp. *croceilyticus*	–	–	Skin
S. petrasii subsp. *petrasii*	–	–	Skin
S. pettenkoferi	–Hotel room	–	Skin
S. piscifermentans	–	Dogs (feces), fermented food and starter cultures	–
S. rostri	–	Pigs, poultry, water buffalo	–
S. saccharolyticus	–	Gorillas	Skin, particularly on the forehead and arm
*S. saprophyticus*subsp. *bovis*	–	Cattle	–
S. saprophyticus subsp. *saprophyticus*		Horses, goats, sheep, cats, fermented food	Skin

Species or subspecies	**Site or source of detection**		
*S. schleiferi*subsp. *schleiferi*	–	Dogs, cats	Skin (particularly preaxillary)
S. sciuri subsp. *carnaticus*	–	Cattle, dolphins	Skin
S. sciuri subsp. *rodentium*	–	Rodents, whales	Skin
S. sciuri subsp. *sciuri*	–Hotel room	Cats, cattle, clams, dogs and other carnivores, dolphins, goats, horses, insectivores, marsupials, monkeys, pigs, rodents, whales	Skin
S. simiae	–	Squirrel monkeys	–
S. simulans	–	Cattle, horses, sheep	Skin (legs, arms, and heads of children; occasionally in adults)
S. stepanovicii	–	Insectivores, rodents	–
*S. succinus*subsp. *casei*	Fermented food	Insectivores, rodents	–

Species or subspecies	Site or source of detection		
*S. succinus*subsp. *succinus*	Amber, fermented food (starter cultures)	Cattle, insectivores, rodents, songbirds	Eye (single report)
S. vitulinus	Fermented food	Horses, poultry	–
S. warneri	Fermented food, Hotel room	Dogs, cats, goats, horses, insectivores, monkeys, pigs, prosimians, rodents, sheep	Skin (preferentially nares, head, legs, and arms)
S. xylosus	Fermented food (starter cultures)	Cats, clams, goats, horses, insectivores, lower primates, rodents, sheep	Skin (rare)

Table 1.3: Sites or source of detection for different staphylococcal species. Table from Becker et al., 2014

1.1.4.1 *S. aureus*

The most well-known *Staphylococcus* spp. is the CoPS *S. aureus.* This microorganism is a human commensal with 30% of the population being an asymptomatic carrier of it. However, it is also one of the most common causes of nosocomial infections. The majority of *S. aureus* carriage is within the anterior nares (Gorwitz et al., 2008). In humans, *S. aureus* can cause bacteraemia, endocarditis, osteoarticular, skin and soft tissue, pleuropulmonary and device-related infections (Tong et al., 2015). It has also been shown to infect pigs, cattle, horses, poultry, sheep, goats and dogs (Peton and Le Loir, 2014). Animal and human-associated isolates *S. aureus* are genetically different though there are cases of animal-associated *Staphylococcus* colonising farmers (Sung et al., 2008). Reports have detected *S. aureus* from environmental samples including beaches, seawater, public buses and built-up area (Conceição et al., 2013; Lutz et al., 2014; Roberts et al., 2013).

1.1.4.2 *S. epidermidis*

S. epidermidis is the most common CoNS human skin-resident; however, rupture of first-line surface barriers such as skin allows the bacteria to gain entry to the bloodstream and become one of the most frequent pathogens among CoNS. It has been evidenced that this bacterium is responsible for 22% of bloodstream infection found in intensive care unit patients in the USA (Otto, 2009). *S. epidermidis* is not exclusively recovered from humans and can be found on animals and plants. For livestock animals, *S. epidermidis* isolates have previously been recovered from cows, pigs and sheep (Argudín et al., 2015; Rahman et al., 2016). Some livestock-associated isolates were shown to be similar to that of hospital-associated isolates though some livestock-associated isolates were only specific to animals (Argudín et al., 2015). Isolates recovered from rice seeds were reported as being genetically different from known human commensal and clinical

isolates but was shown to be genetically similar to isolates recovered from wild mice (Chaudhry and Patil, 2016). MDR *S. epidermidis* has been isolated from non-healthcare settings like hotel rooms and university campuses, but no studies have shown if they are similar to that of hospital or animal associated *S. epidermidis* isolates (Seng et al., 2017b; Xu et al., 2015).

1.1.4.3 *S. haemolyticus*

S. haemolyticus is the second most common CoNS species that have been isolated from clinical cases (Czekaj et al., 2015). Similar to *S. epidermidis*, most infections are associated with immune-compromised patients and patients with implanted medical devices (Silva et al., 2013). *S. haemolyticus* has the highest level of antimicrobial resistance among the CoNS (Barros et al., 2012; Froggatt et al., 1989). Reports have shown that *S. haemolyticus* was found in 38.3% of the infant's nasal cavity from 429 neonates admitted in a hospital in Brazil (Ternes et al., 2013). They have also been cases of AMR *S. haemolyticus* recovered from livestock, companion animals and public environments (Alirol et al., 2017; Ruzauskas et al., 2014; Seng et al., 2017b; Xu et al., 2015). Unlike *S. epidermidis* little is known about genetic lineages of isolates recovered from a different source, but studies have reported on the transmission of isolates from companion and livestock animals to humans (Loncaric et al., 2016; Ruzauskas et al., 2014).

1.1.4.4 *S. hominis*

S. hominis like *S. epidermidis* and *S. haemolyticus* is another CoNS human commensal, which has also been associated with nosocomial (Chaves et al., 2005; Ternes et al., 2013). Reports of *S. hominis* colonisation on animals is scarce, with only a few reports of them being isolated from dairy cattle and mosquitos (Hughes et al., 2016; Naushad et al., 2016). These reports do not show if these isolates could be transmitted to humans. From the

public settings, *S. hominis* has been previously isolated from hotel rooms as well as university campuses that were shown to be MDR (Seng et al., 2017b; Xu et al., 2015). No previous studies have looked into the genetic lineages of animal or environmental *S. hominis* before.

1.2 Pathogenicity

Staphylococci are capable of colonising and infecting many different hosts. These virulent factors can be species and/or strain-specific that are shown to have a function in adherence, aggression, invasion, persistence and evasion of the adaptive and innate immune system (Table 1.3) (Diep et al., 2006; Gill et al., 2005). Of the virulent factors previously characterised for this genus, the majority of them have only been identified in *S. aureus* (Gill et al., 2005). This includes a wide range of toxin genes which can cause diseases such as toxic shock syndrome (TSS), staphylococcal scalded skin syndrome (SSSS), necrotizing pneumonia, or deep-seated infections (Jarraud et al., 2002, 1999). In comparison, CoNS lack many of the virulent factors that have been identified in *S. aureus*. However, CoNS are still able to cause infections, although not to the degree to that of *S. aureus*. CoNS infections are mainly caused by the implantation of medical devices (Christensen et al., 1985; Peters et al., 1982). Attachment to these devices normally requires the bacteria to produce a biofilm. Infections normally occur during the insertion of the device as a small number of cells attach to the implant from the patient's skin or mucosal layer. From here, the cells can disseminate via the bloodstream and colonise and infect other body sites, which can lead to sepsis, meningitis and endocarditis. Another important virulent factor that has been identified in *S. haemolyticus*, *S. hominis*, *S. saprophyticus* as well as *S. aureus* is the production of a polysaccharide capsule (Blake and Metcalfe, 2001; Flahaut et al., 2008; O'Riordan and Lee, 2004; Park et al., 2010). This capsule is important in host immune invasion (Cunnion et al., 2003).

Adherence	Gene name	*S. aureus*	*S. epidermidis*	*S. haemolyticus*
Autolysin	*atl*	P	p	P
Cell wall-associated fibronectin binding protein	*ebh*	P	p	A
Clumping factor A	*clfA*	P	A	A
Clumping factor B	*clfB*	P	A	A
Collagen adhesion	*cna*	P	A	A
Elastin binding protein	*ebp*	P	p	p
Extracellular adherence protein/MHC analogous protein	*eap/map*	P	A	A
Fibrinogen binding protein	*efb*	P	A	A
Fibronectin binding proteins	*fnbA*	P	A	A
Fibronectin binding proteins	*fnbB*	P	A	A
Intercellular adhesin	*icaA*	P	P	A
Intercellular adhesin	*icaB*	P	P	A
Intercellular adhesin	*icaC*	P	P	A
Intercellular adhesin	*icaD*	P	P	A
Intercellular adhesin	*icaR*	P	P	A
Ser-Asp rich fibrinogen-binding proteins	*sdrC*	P	A	A
Ser-Asp rich fibrinogen-binding proteins	*sdrD*	P	A	A
Ser-Asp rich fibrinogen-binding proteins	*sdrE*	P	A	A
Ser-Asp rich fibrinogen-binding proteins	*sdrF*	A	P	A
Ser-Asp rich fibrinogen-binding proteins	*sdrG*	A	P	A
Ser-Asp rich fibrinogen-binding proteins	*sdrH*	A	P	A
Staphylococcal protein A	*spa*	P	A	A
Enzyme				
Cysteine protease	*sspB*	P	P	A
Cysteine protease	*sspC*	P	A	A
Hyaluronate lyase	*hysA*	P	A	A
Lipase	*geh*	P	P	A
Lipase	*lip*	P	P	p
Serine V8 protease	*sspA*	P	P	A

Adherence	Gene name	*S. aureus*	*S. epidermidis*	*S. haemolyticus*
Serine protease	*splA*	P	A	A
Serine protease	*splB*	P	A	A
Serine protease	*splC*	P	A	A
Serine protease	*splD*	P	A	A
Serine protease	*splE*	P	A	A
Serine protease	*splF*	P	A	A
Staphylocoagulase	*coa*	P	A	A
Staphylokinase	*sak*	P	A	A
Thermonuclease	*nuc*	P	P	P
Immune evasion				
Capsule	--	P	A	P
Secretion system				
	esxA			
Type VII secretion system	*esaA*	P	A	A
Type VII secretion system	*essA*	P	A	A
Type VII secretion system	*esaB*	P	A	A
Type VII secretion system	*essB*	P	A	A
Type VII secretion system	*essC*	P	A	A
Type VII secretion system	*esxB*	P	A	A
Type VII secretion system	*esaD*	P	A	A
Type VII secretion system	*esaE*	P	A	A
Type VII secretion system	*esxD*	P	A	A
Type VII secretion system	*esaD*	P	A	A
Type VII secretion system	*esaG*	P	A	A
Toxin				
Alpha hemolysin	*hly/hla*	P	A	A
Beta hemolysin	*hlb*	P	P	A
Delta hemolysin	*hld*	P	P	A
Enterotoxin A	*sea*	P	A	A
Enterotoxin B	*seb*	P	A	A
Enterotoxin C	*sec*	P	A	A
Enterotoxin D	*sed*	P	A	A

Adherence	Gene name	*S. aureus*	*S. epidermidis*	*S. haemolyticus*
Enterotoxin E	*see*	P	A	A
Enterotoxin G	*seg*	P	A	A
Enterotoxin H	*seh*	P	A	A
Enterotoxin I	*sei*	P	A	A
Enterotoxin J	*sej*	P	A	A
Enterotoxin Yent1	*yent1*	P	A	A
Enterotoxin Yent2	*yent2*	P	A	A
Enterotoxin-like K	*selk*	P	A	A
Enterotoxin-like L	*sell*	P	A	A
Enterotoxin-like M	*selm*	P	A	A
Enterotoxin-like N	*seln*	P	A	A
Enterotoxin-like O	*selo*	P	A	A
Enterotoxin-like P	*selp*	P	A	A
Enterotoxin-like Q	*selq*	P	A	A
Enterotoxin-like R	*selr*	P	A	A
Enterotoxin-like U	*selu*	P	A	A
Exfoliative toxin type A	*eta*	P	A	A
Exfoliative toxin type B	*etb*	P	A	A
Exfoliative toxin type C	*etc*	P	A	A
Exfoliative toxin type D	*etd*	P	A	A
Exotoxin	*set10*	P	A	A
Exotoxin	*set11*	P	A	A
Exotoxin	*set12*	P	A	A
Exotoxin	*set13*	P	A	A
Exotoxin	*set14*	P	A	A
Exotoxin	*set15*	P	A	A
Exotoxin	*set16*	P	A	A
Exotoxin	*set17*	P	A	A
Exotoxin	*set18*	P	A	A
Exotoxin	*set19*	P	A	A
Exotoxin	*set1*	P	A	A
Exotoxin	*set20*	P	A	A
Exotoxin	*set21*	P	A	A
Exotoxin	*set22*	P	A	A
Exotoxin	*set23*	P	A	A
Exotoxin	*set24*	P	A	A
Exotoxin	*set25*	P	A	A
Exotoxin	*set26*	P	A	A
Exotoxin	*set2*	P	A	A
Exotoxin	*set30*	P	A	A

Adherence	Gene name	*S. aureus*	*S. epidermidis*	*S. haemolyticus*
Exotoxin	*set31*	P	A	A
Exotoxin	*set32*	P	A	A
Exotoxin	*set33*	P	A	A
Exotoxin	*set34*	P	A	A
Exotoxin	*set35*	P	A	A
Exotoxin	*set36*	P	A	A
Exotoxin	*set37*	P	A	A
Exotoxin	*set38*	P	A	A
Exotoxin	*set39*	P	A	A
Exotoxin	*set3*	P	A	A
Exotoxin	*set40*	P	A	A
Exotoxin	*set4*	P	A	A
Exotoxin	*set5*	P	A	A
Exotoxin	*set6*	P	A	A
Exotoxin	*set7*	P	A	A
Exotoxin	*set8*	P	A	A
Exotoxin	*set9*	P	A	A
Gamma hemolysin	*hlgA*	P	A	A
Gamma hemolysin	*hlgB*	P	A	A
Gamma hemolysin	*hlgC*	P	A	A
Leukocidin M	*lukF-like*	P	A	A
Leukocidin M	*lukM*	P	A	A
Leukotoxin D	*lukD*	P	A	A
Leukotoxin E	*lukE*	P	A	A
Panton-Valentine leukocidin	*lukF-PV*	P	A	A
Panton-Valentine leukocidin	*lukS-PV*	P	A	A
Toxic shock syndrome toxin	*tsst*	P	A	A

Table 1.4: Known virulent genes found in different staphylococcal species. Part of table data from http://www.mgc.ac.cn/cgi-bin/VFs/v5/main.cgi?func=VFanalyzer Staphylococcus pathogenomics accessed March 2018.P= present A= absent

1.3 Speciation of staphylococci

Speciation of staphylococci has developed from a long process of bacteria cultivation and biochemical tests for rapid identification. The improvement in speciation is important in clinical laboratories to help rapidly identify bacterial species cost-effectively to ensure that patients are given the right treatment quickly.

1.3.1 Traditional biochemical speciation methods

A traditional method used in identifying staphylococci is to grow them on selective or differential media. A common media used in the identification of staphylococci is mannitol salt agar (Bautista-Trujillo et al., 2013). This agar is selective to staphylococci due to the high sodium chloride concentration (7.5%) and can differentiate between *S. aureus* and other staphylococci by the pH indicator which changes the media from red to yellow due to the fermentation of mannitol into an acid by-product by *S. aureus* (Bautista-Trujillo et al., 2013). Although this might be a quick method to identify *S. aureus* there are instances of *S. aureus* that are mannitol-negative as well as some *S. saprophyticus* and *S. haemolyticus* being mannitol positive (dos Santos et al., 2015; Shittu et al., 2006). This agar does not exclusively selective for staphylococci therefore further analyse would be required. To determine if these might be *Staphylococcus* and not bacteria from another genus; Gram staining would be performed followed by a series of biochemical tests (Figure 1.1). These would include a catalase test to determine if they might be staphylococci and a coagulase test which will determine if the bacteria are CoPS or CoNS (Figure 1.2). This method for speciating bacteria is time consuming and can normally take 24-48 hours to complete (Croxatto et al., 2012) .

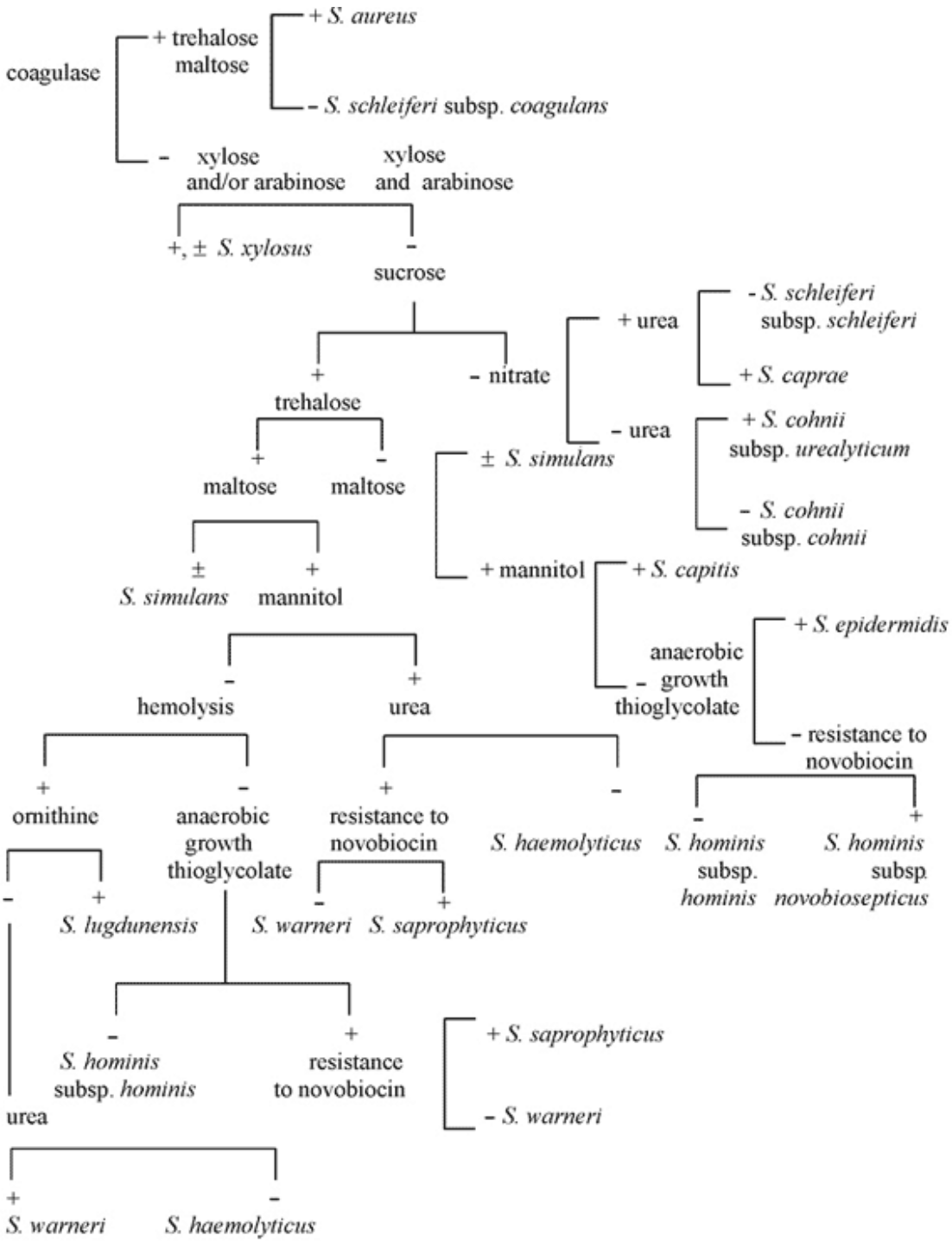
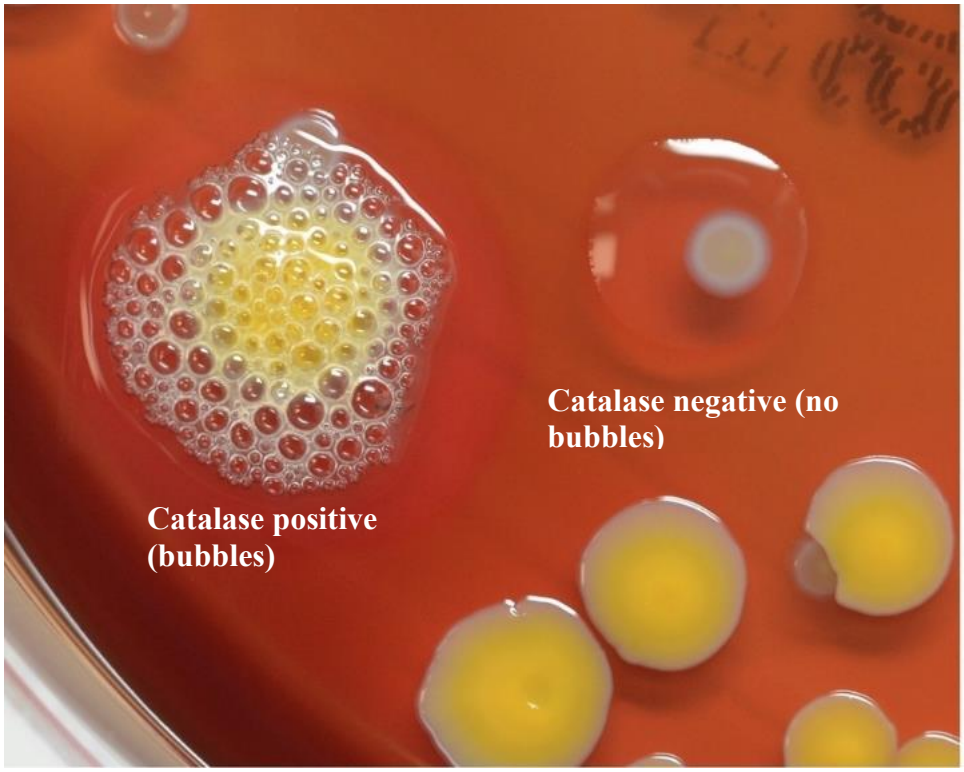


Figure 1.1: **Schematic of biochemical test to speciation of staphylococci. Figure from https://bio.libretexts.org/Ancillary_Materials/Experiments/Microbiology_Labs_I/22A%3A_Identification_of_*Staphylococcus*_species**

A



B

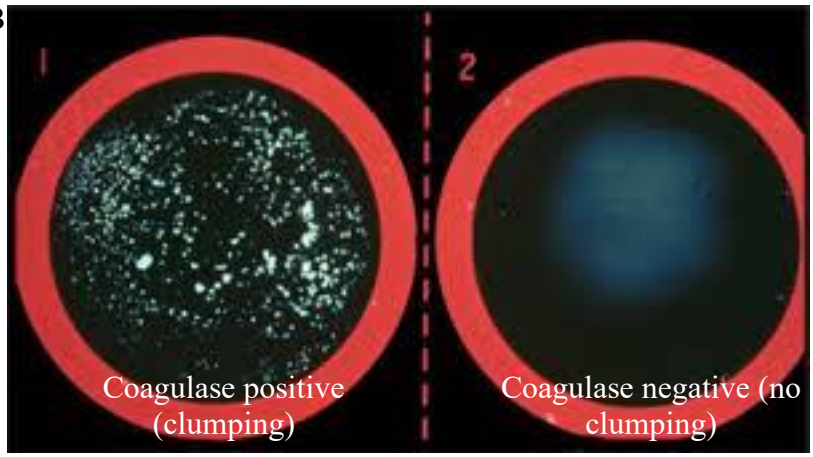


Figure 1.2: Catalase and coagulase test for staphylococcal identification. (A) Catalase test figure https://www.microbiologyinpictures.com/bacteria-photos/staphylococcus-aureus-photos/s_aureus_tests.html (B) Coagulase test figure from https://dph.georgia.gov/sites/dph.georgia.gov/files/related_files/site_page/The%20Impact%20of%20Rapid%20Diagnostics%20on%20Antimicrobial%20Stewardship.pdf

1.3.2 16s ribosomal RNA gene sequencing

A modern method used in t speciation of staphylococci is 16S ribosomal RNA (rRNA) gene sequencing (Janda and Abbott, 2007). 16S rRNA gene is a component of the 30S subunit of prokaryotic ribosomes. This gene is highly conserved between different species and can be amplified by PCR using universal primers to produce a 1,500bp sequence (Weisburg et al., 1991). Using bioinformatics tools, the isolates DNA sequence are then compared to known sequences on an open database to identify the bacterial species, and sometimes to strain level (Janda and Abbott, 2007). This method only works as long as that sequence has been identified before (Janda and Abbott, 2007). Compared to biochemical test this method is quicker and more accurate but still relatively time consuming.

1.3.3 Speciation by matrix-assisted laser desorption ionization time-of-flight mass spectrometry

Recently, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) has become widely used in clinical microbiology for rapid identification of bacteria to the species level (Croxatto et al., 2012) (Figure 1.3). This method is cheaper and faster of that used to in 16s rRNA but is not as efficient in identifying the different strains (Croxatto et al., 2012). This method only requires a single colony from an agar plate to be put onto a target plate. The cells on the target plate are then lysed with acid before a matrix solution is overlaid. They are then put into the MALDI-TOF MS instrument, and a laser is fired multiple times onto each spot, which leads to desorption of the analyte which are then vaporised and ionised. These desorbed and ionised molecules are first accelerated through an electrostatic field then through a vacuum until they reach a detector, with small ions travelling quicker than larger ions.

This detector measures intensity and mass/charge (m/z) of each analyte fragment and plots them on a spectrum. The spectrum of these samples is then compared with a database of known bacterial species key mass ions to identify the bacterial species (Croxatto et al., 2012). A drawback with this method is that there are several manufacturers that produce MALDI-TOF MS each having their own database has restricted its commercial development, currently, to two major companies (Singhal et al., 2015). Parallel studies with clinical isolates using Bruker's Biotyper and the Shimadzu's MALDI-MS using the SARAMIS system of bioMérieux has shown good concordance when reference spectra are present in the database (Cherkaoui et al., 2010). However, while the degree of confidence of a result mainly depends on the MALDI-MS and database, it is also influenced by sample preparation and even user dependent (Keys et al., 2004).

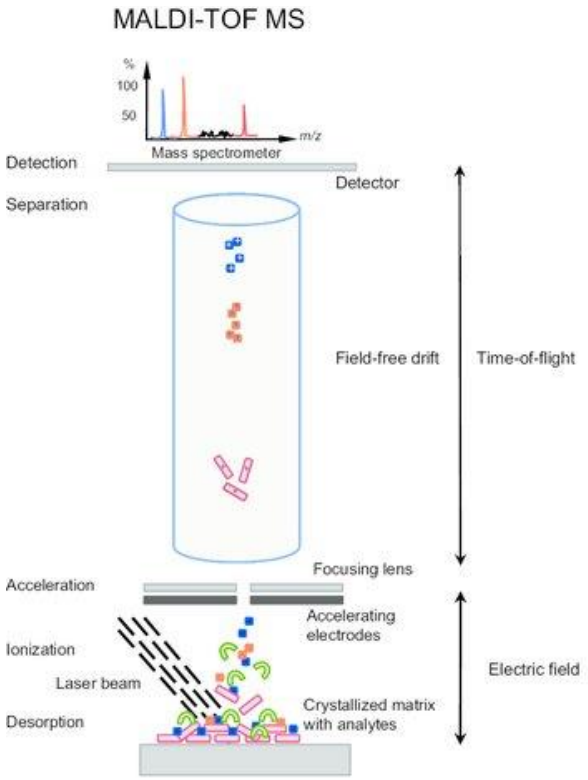


Figure 1.3: Principal of MALDI-TOF identification of bacterial species. Figure from Lavigne et al., 2013

1.4 Molecular typing of staphylococci

Molecular typing is important for monitoring the staphylococci population as well as identifying if the infections are related to an outbreak (Miragaia et al., 2008). As bacteria species are constantly diversifying due to point mutation, recombination, acquiring or deleting of mobile genetic elements, it not always easy to find genetic markers that can be traced back to a common ancestor. Therefore genetic markers that are used for typing bacteria isolates need to be reproducible, stable; have high discriminatory power, epidemiological concordance versatile, easy to perform, easy to interpret; and should be cost-effective and time effective (Miragaia et al., 2008).

1.4.1 Pulse field gel electrophoresis

A traditional method used to group variants in many bacterial species is pulse field gel electrophoresis (PFGE) (Miragaia et al., 2008). This method uses nucleotide restriction enzymes to digest chromosomal DNA into short nucleotide fragments. The banding pattern of the fragments is then compared on an electrophoresis gel (Oliveira et al., 2002).

1.4.2 Multilocus sequence typing

Some bacteria species can also be grouped by Multilocus sequence typing (MLST) (Figure 1.4). This requires amplifying fragments of seven housekeeping genes of individual species and sequencing them to identify polymorphism(s) in those genes (Miragaia et al., 2008). These strains are then grouped by their nucleotide polymorphisms into sequence type (ST). ST can be further grouped by their clonal complex (CC), which are ST distinguishable by a single nucleotide polymorphism in three or fewer loci

(Chambers and Deleo, 2009). MLST in *S. aureus* has been used to distinguish between hospital-associated and community-associated infection as well as to track the movement of different outbreaks across the world (Ghaznavi-Rad et al., 2010; Green et al., 2010; Wang et al., 2018).

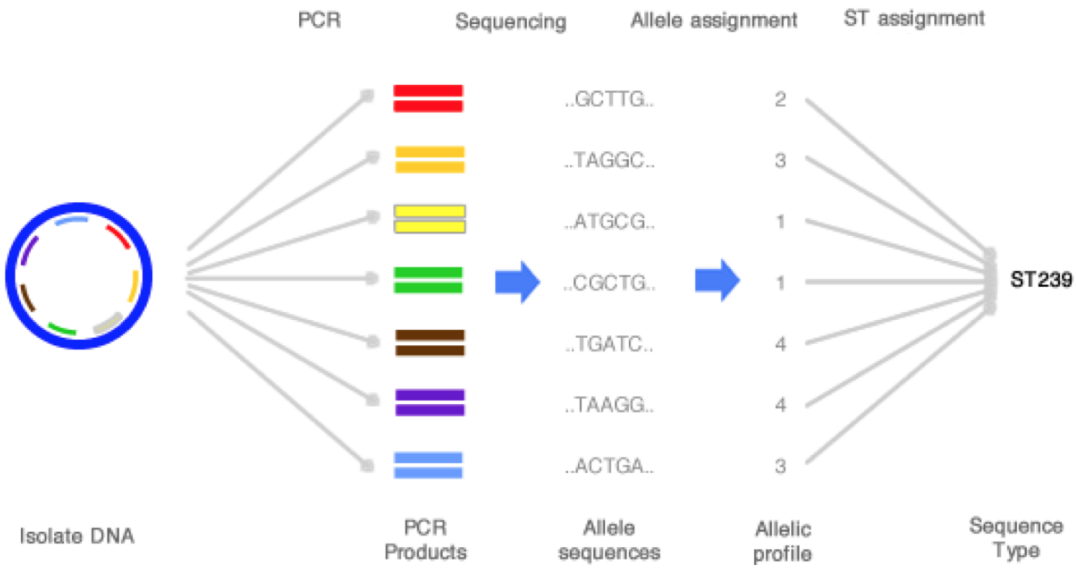


Figure 1.4: Schematic on how MLST is performed. Seven housekeeping genes are amplified from isolate and sequenced to determine their allele type. From the allelic profile the bacteria isolate ST can be determined. Figure from (**http://beta.mlst.net/)**

1.4.3 SCC*mec* typing

Staphylococcal isolates can also be grouped by their *Staphylococcus* Chromosome Cassette *mec* (SCC*mec*) type a mobile genetic element that has the *mecA* gene which encodes for methicillin resistance (Miragaia et al., 2008). This method helps determine the movement of these genes in different populations as well as to understand the SCC*mec* evolution (Miragaia et al., 2008). SCC*mec* typing can only be used in strains that have the genetic element where MLST uses genes that are ubiquitous in *S. aureus* and *S. epidermidis*.

1.4.4 *Spa* typing

The *Spa* gene variable repeat region can also be used to differentiate between *S. aureus* isolates. *Spa* gene encodes protein A, which is an important virulence factor in *S. aureus* (Koreen et al., 2004). In *Spa* typing, they can discriminate between different *S. aureus* in outbreak settings by sequencing the polymorphic 24-bp variable-number tandem repeat (VNTR) within the 3′ coding region. Genetic variations within the spa gene can occur rapidly and slowly by two independent mechanisms which can be used as a marker to track local and global transmission as well as long term epidemiologic and population studies (Koreen et al., 2004).

1.4.5 Comparison of different molecular typing methods

Different molecular typing techniques have their strengths and weaknesses. This was shown in two studies performed by Miragaia and co-workers, and Petersson and co-workers (Miragaia et al. 2008; Petersson et al., 2010). These studies compared PFGE, MLST and SCC*mec* in *S. epidermidis* and *Spa* type to PFGE in *S. aureus.* The first study found that PFGE is most discriminative, followed by MLST and then SCC*mec* (Miragaia et al., 2008). The second study found that *Spa* was less discriminative in comparison to PFGE (Petersson et al., 2010). Although MLST might not be as discriminative as PFGE as shown by Miragaia and co-workers; MLST is used more today than PFGE as the techniques are standardised and results can be compared with other laboratories isolates submitted into the database, which is freely available (Miragaia et al. 2008; Nemoy et al., 2005). *Spa* typing might not be as discriminating as the other techniques but still can be useful in epidemiology cases in a low-prevalence setting (Petersson et al., 2010).

1.5 Whole genome sequencing

The most discriminative method used today in determining the difference in genetics in a bacterial population is whole genome sequencing (WGS) (Bryant et al., 2012). This method generates a sequence of the whole bacterial chromosomes and plasmids for each species. WGS has helped aid in the understanding of the epidemiology, genetic evolution between strains and species and the frequency these genes occur in that species (Aanensen et al., 2016; Conlan et al., 2012; Gill et al., 2005). In the past WGS could not have been done on a large scale due to cost and time but advances in next-generation sequencing in the last 10 years has helped significantly by decreasing these limiting factors (Goodwin et al., 2016). This coincides with the advancement of computers ability to process a large amount of biological data quickly have shown an increase in studies using WGS and bioinformatics analyses in microbiology (Saeb, 2018). However, most of the literature still uses standard molecular typing for staphylococci as it still is considered quicker for routine surveillance and there is still some uncertainty on how to interpret the WGS data for epidemiology (Sabat et al., 2013).

1.5.1 WGS to compare different species

WGS can be used to compare many different bacteria species. An example of this for staphylococci is a study that compared *S. aureus* COL; one of the first methicillin-resistant isolates and methicillin-resistant *S. epidermidis* RP62a (Gill et al., 2005). They found that these two species shared 1,681 core genes and the only variation most likely derives from genome islands found on the bacterial chromosome. Interestingly, the *S. epidermidis* strain they sequenced obtained a unique virulence gene for the species which encodes a polyglutamate capsule similar to that found in *Bacillus anthracis* (Gill et al., 2005). This shows that WGS can be used to identify whether genes have been horizontally transferred from different bacterial species.

1.5.2 WGS to compare isolates from same species

Most WGS studies are usually comparing populations of a single bacterial species. An example of this for staphylococci is a study which compared an environmental *S. epidermis* G6_2 isolate recovered from a hotel room in London to previous WGS *S. epidermidis* isolates (Xu et al., 2018a). In this study, they identified this isolate to be a representative of MLST 59 (Xu et al., 2018a). They determined that this isolate was similar to that of other ST59 isolates phylogenetically and that the composition of the antibiotic resistance gene is partly preserved across this lineage (Xu et al., 2018a). Interestingly, the virulence delta-hemolysin gene was present in *S. epidermis* G6_2, which was absent in other ST59 samples (Xu et al., 2018a).

WGS can be used to analyse the bacterial population structure. A common method of analysing bacterial population structure is by a pangenome analysis of bacterial species by considering which genes are considered core and which genes are accessory within a species (Conlan et al., 2012). Conlan and co-workers WGS to identify that *S. epidermidis* that are known as commensal and nosocomial (Conlan et al., 2012). They found that *S. epidermidis* species have a large core genome, but variable genes are constantly in flux due to transposable elements transcription factors and transporters (Conlan et al. 2012). From just WGS data, they could distinguish between commensal and nosocomial strains by the reduction of virulence markers and the presence of formate dehydrogenase gene (Conlan et al., 2012). They also found that nosocomial strains were shown to have more genetic rearrangement and single nucleotide polymorphism, than the commensal strains (Conlan et al., 2012). Additionally, they identified gene functions within the variable genome by assigning gene clusters of orthologous groups (COG) categories by aligning protein sequence against a BLAST database of COG sequences. This consists of 20 groups, of which they found that there was high abundance of genes in the variable

genome which encodes for replication, recombination and repair; transcriptional regulators; and defense mechanisms (Conlan et al., 2012; Tatusov et al., 2003). The authors describe the reasoning for the high abundance of these COG groups in the variable genome was due high diversity of mobile genetic elements (recombinase and integrase genes), transcriptional regulators, and ABC-type multidrug transporters, respectively (Conlan et al., 2012).

1.5.3 Detecting mobile genetic elements

A key factor of the antibiotic resistance genes and virulence genes is that they can be found within mobile genetic elements. From WGS data, there are many approaches which have been implemented in trying to identifying if these genes are horizontally transferred across (Lu and Leong, 2016). These approaches including using NCBI BLAST for known mobile genetic elements found in phage insertion sequences or transposons and insertion sequences; genome assembling software that detect plasmids by comparing the read coverage to the overall medium coverage of the genome and software that can detect sequence composition biases to determine areas within the genome are genomic islands (Altschul et al., 1990; Antipov et al., 2016; Arndt et al., 2016; Bertelli et al., 2017). It is also possible to identify the donor's organisms of the horizontal transfer genes by blasting proteins sequences against the NCBI reference database of proteins with known species (Pruitt et al., 2007; Zhu et al., 2014).

Staphylococci WGS analyses other than *S. aureus* and *S. epidermidis* are very fragmented in the literature (Cavanagh et al., 2014; Conlan et al., 2012; Gill et al., 2005; Lilje et al., 2017; Sabat et al., 2013; Xu et al., 2015). This may because other staphylococci are not common causes of infection. This might be the case but there has been an increasing number of *S. haemolyticus* and *S. hominis* infections that have been reported and it would

be interesting to find out more about their genetic lineages, especially the ones which are MDR (Czekaj et al., 2015; Voineagu et al., 2012).

1.6 Antibiotic resistance variations in isolates recovered from different ecological niches

MDR staphylococci have been found in many different ecological niches; ranging from healthcare, community, and environmental sources. The majority of the literature focuses on *S. aureus* isolates from the clinical setting (patients and environment) and the community (isolates from people not associated with healthcare facilities) but reports are more fragmented on the public environment especially for CoNS.

1.6.1 MRSA population

One of the most studied bacterial populations for staphylococci is methicillin-resistant *S. aureus* (MRSA). MRSA is interesting as it can be categorised into 3 subgroups, healthcare-associated MRSA (HA-MRSA); community-associated MRSA (CA-MRSA) and livestock-associated MRSA (LS-MRSA) (Naimi, 2003; Vanderhaeghen et al., 2010). All three groups have been shown to evolve separately from different clonal backgrounds (Naimi, 2003; Vanderhaeghen et al., 2010). HA-MRSA isolates are transferred in healthcare facilities either from patients or healthcare workers; CA-MRSA is isolated from people who have a minimum to no contact with healthcare facilities whereas LA-MRSA isolates only affect animals normally used in agriculture (Naimi, 2003; Vanderhaeghen et al., 2010). These groups were shown to have different phenotypes, for example, HA-MRSA was more resistant to antibiotics whereas CA-MRSA was more virulent due to them carrying the Panton-Valentine Leukocidin (PVL) toxin (Naimi, 2003). It has also been shown that SCC*mec* type I, II and III are associated with HA-MRSA whereas the other SCC*mec* types are associated with CA-MRSA and LA-MRSA

(Ahmad et al., 2009; Köck et al., 2013). A study in a hospital in Chicago found that HA-MRSA (in epidemiological terms) isolates are more phenotypically related to what is consider community-associated based on their virulence and SCC*mec* element type (Maree et al., 2007). For LA-MRSA, there are few reports of human infections (Becker et al., 2017; Dorado-García et al., 2013)

1.6.2 MRSA in public settings

HA-MRSA isolates have been recently reported to be recovered from public settings not normally associated with healthcare facilities. These reports include isolates recovered from public beach sands, fresh and marine water; university campus and on public buses (Akanbi et al., 2017; Conceição et al., 2013; Lutz et al., 2014; Roberts et al., 2013; Soge et al., 2009). Studies on public buses in Portugal and the USA both reports that high-frequency touched surfaces had a mixture of HA-MRSA and CA-MRSA isolate (Conceição et al., 2013; Lutz et al., 2014). The transfer of HA-MRSA onto public buses surfaces was due to bus routes going towards hospitals, therefore, were transmitted by patients and hospital workers. For public beaches reports from both the USA and South Africa were found to have CA-MRSA and HA-MRSA in beach sand and marine water (Akanbi et al., 2017; Roberts et al., 2013; Soge et al., 2009). The exact route of transfer of HA-MRSA is unknown, but it was discussed that higher prevalence of HA-MRSA in marine water than beach sand is due to higher exposure of contamination from pharmaceuticals, hospitals, and industrial waste as well as farmlands (Akanbi et al., 2017). From high-frequency touched surfaces in the public area on a university campus in the USA, there were MRSA clones that were similar to healthcare and community-associated origins (Roberts et al., 2013). In this study, the transmission of HA-MRSA to high-frequency touched surfaces is unknown (Roberts et al., 2013). Interestingly, there are no reports in the literature on LA-MRSA isolates found in the public setting.

1.6.3 Healthcare and community-associated CoNS

There are no reports on CoNS methicillin-resistant isolates distinction of their genotype and phenotype by isolation source but there is a study that identified the potential settings they are more likely to be spread. This study was published in 2010 in which they used long-range PCR to detect SCC*mec* element from nasal carriage CoNS isolates from patients who had no previous exposure to hospitals (Barbier et al., 2010). In this study, they tested 291 patients and found that 56 of them had methicillin resistance CoNS (Barbier et al., 2010). The species found in this study were methicillin resistant was *S. epidermidis*, *S. hominis*, *S. haemolyticus*, *S. pettenkoferi, S. cohnii* (Barbier et al., 2010). This suggested that these species of staphylococci were more likely to spread in a community than the hospital, especially in patients who had no previous exposure to hospitals (Barbier et al., 2010).

1.6.4 CoNS in public settings

There are a few studies which have looked into the population of CoNS in public settings. Xu and co-workers reported that there were 71 isolates from 11 CoNS species from hotels rooms in London which were MDR as well as some of them showing the presence of *mecA* gene (Xu et al., 2015). In this study, the authors used PCR to identify the SCC*mec* and MLST (Xu et al., 2015). The most commonly detected species they found in this environment were *S. haemolyticus*, *S. hominis, S. warneri, S. cohnii*, and *S. epidermidis* (Xu et al., 2015). Interestingly, they also reported new *S. epidermidis* MLST types (Xu et al., 2015). Some of the isolates which had the *mecA* gene had a lower level of methicillin resistance to that which was expected as we; as some of the isolates had *SCCmec* type that was undiscovered (Xu et al., 2015). This suggests that these environments could be reservoirs for MDR staphylococci (Xu et al., 2015). Seng and co-

workers reportedly found that 41 out of 200 (20.5%) CoNS samples isolated from the university campus in Thailand were methicillin resistant (Seng et al., 2017b). These belonged to 6 different species, including *S. haemolyticus*, *S. hominis, S. warneri, S. cohnii*, *S. epidermidis and S. saprophyticus* (Seng et al., 2017b). In a similar manner to Xu and co-workers they used PCR to identify the SCC*mec* type and found that they belong to multiple types but the majority of these isolates SCC*mec* were untypable (Xu et al., 2015, Seng et al., 2017b).

1.6.5 Bacterial “resistomes"

Bacterial "resistomes" is the study of all the genes that confer antibiotic resistance and their precursor in bacteria in a particular microbiome. As described by Wright it includes antibiotic resistance genes from pathogens, antibiotic resistance genes in antibiotic-producing bacteria (to protect themselves from antibiotics), cryptic resistance genes (genes found on the chromosome that would not obviously confer resistance due to low or lack of expression) and precursor genes (genes that do not directly confer resistances but may encode a protein that may have a level of activity against antibiotic molecule or have affinity to the molecule in which, depending on the right selective pressure, may evolve to a full resistant gene) (Wright, 2007). Resistome studies aims are to better understand the spread of antibiotic resistance genes from different environments as well as to understand the origins of antibiotic resistance genes. One resistome study that focused on staphylococci from 18 public restrooms from 4 buildings from non-healthcare settings over period of 24 weeks found that many of the staphylococcal species isolated were resistant to antibiotics have similar antibiograms to different species from the same restroom on different dates and others with isolates from restrooms in the same building (Mkrtchyan et al., 2013). They were able to identify the direct transfer of resistance determinants within restrooms and/or within buildings as they found that 11

staphylococcal isolates with the same antibiograms representing 5 different staphylococcal species and these were isolated from 5 different restrooms within the same building (Mkrtchyan et al., 2013). Additionally, they identified that from a single restroom there were 6 isolates from 3 species (*S. haemolyticus* (n=3), *S. epidermidis* (n=2) and *S. aureus*) from different sites that had the same antibiogram. These antibiograms were demonstrated to be also present in other staphylococcal species in 4 other restrooms within the same building on different days indicating widespread dissemination of resistance determinants in different *Staphylococcus* species and restrooms in the same building. They were able to demonstrate that public restrooms "resistome" and that these areas have a collection of pathogenic and non-pathogenic bacteria with antibiotic resistance determinants (Mkrtchyan et al., 2013). Other studies in "resistome" have focused on all bacterial species within a microbiome using metagenomic analyses in soil, hospitals wastewater, community sewage and gut microbiome then comparing a resistome within a single genus (Buelow et al., 2018; Mafiz et al., 2018; van Schaik, 2015).

More research is required to understand multidrug resistance staphylococcal isolates recovered from the environment as different reservoirs may have different antibiotic resistance profiles; SCC*mec* types, new species genotypes and virulent factors previously undiscovered in hospital-associated, community-associated and livestock-associated strains. The majority of work done on environmental isolates uses standard molecular typing but does not reveal the whole picture (Conceição et al., 2013; Lutz et al., 2014; Seng et al., 2017b, Xu et al., 2015). Using WGS will show if their genetic background is similar to that of hospital, community or animal associated staphylococcal isolates. It will also reveal whether any novel genes found in only these isolates are acquired from other bacteria or similar lineages. Using the "One Health" approach which is an ideology of

comparing AMR bacteria in humans, animals and their environment, to determine staphylococcal isolates from public settings are a public health risk (Walsh, 2018).

1.7 Antibiotic resistance

The mass production of many types of antibiotics has revolutionised modern medicine; drastically reducing the mortality caused by bacterial infections. However, more and more bacterial infections are no longer treatable with modern antibiotics with a possibility of a future where mortality from bacterial infection will be high (Hau et al., 2018). AMR is not a modern phenomenon and precedes that of the first mass-produced antibiotics (D'Costa et al., 2011). This is due to the fact that antibiotics can be found naturally in fungus and bacteria as a defence mechanism (D'Costa et al., 2011). It only became a serious problem when antibiotics were overused in healthcare and agriculture (Shallcross and Davies, 2014). The development of AMR can arise through horizontal gene transfer by species to species; bacteriophages, plasmid or transposons pathogenicity islands, chromosomal cassettes as well as random mutations driven by selection when in the presence of antibiotics or other stress factors from its surrounding environment (Fajardo et al., 2008; Malachowa and DeLeo, 2010; Martinez and Baquero, 2000; Munita and Arias, 2016)

Staphylococci have been found to harbour many different antibiotic resistance genes that give resistance to many types of drug classes which can be found in bacteria population associated with hospitals; community; animal and isolates that are found in the environment (Table 1.4) (Ahmad et al., 2009; Conceição et al., 2013; Soge et al., 2009; Xu et al., 2015). In the past hospital and agriculture isolates were shown to have resistance to many types of antibiotics compared to isolates from other sources due to the higher abundance of antibiotics. A recent study from Taiwan noticed little difference in the antibiotic resistance profile in isolates that cause keratitis from HA-MRSA and CA-

MRSA whereas another study was shown to have a significantly higher portion of CA-MRSA that were resistant to ciprofloxacin and clindamycin compared to HA-MRSA (Hsiao et al., 2015; Huang et al., 2006). In a study that looked at ST5 MRSA from human clinical samples and samples taken from swine in the USA were shown to be MDR, however, the two groups were distinct based on their phenotype and genotype (Hau et al., 2018). There were no reports comparing the difference between antibiotic resistance profile between CoNS hospital-associated infection and community infection and livestock though studies have shown that both areas isolates showed resistance to multiple antibiotics with a rise of MDR isolates over time found in hospital and the community associated isolates (Bhargava and Zhang, 2012; May et al., 2014; Nanoukon et al., 2017) . Studies have found MDR resistance S. *aureus* and CoNS have been detected from the environment including within hospital environments, home environment, hotel rooms, university campuses and livestock environments (Schoenfelder et al., 2017; Seng et al., 2017b; Shahbazian et al., 2017; Xu et al., 2015).

Gene	Resistant drug group and function	*S. aureus*	*S. epidermidis*
AAC(6')-Ib7	aminoglycoside antibiotic; antibiotic inactivation	P	A
AAC(6')-Ie-APH(2'')-Ia	aminoglycoside antibiotic; antibiotic inactivation	P	P
AAC(6')-Ig	aminoglycoside antibiotic; antibiotic inactivation	P	P
aad(6)	aminoglycoside antibiotic; antibiotic inactivation	P	P
aadA	aminoglycoside antibiotic; antibiotic inactivation	P	A
acrB	antibiotic efflux; cephalosporin; fluoroquinolone antibiotic; glycylcycline; penam; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic; triclosan	P	A
acrD	aminoglycoside antibiotic; antibiotic efflux	P	P
acrF	antibiotic efflux; cephalosporin; cephamycin; fluoroquinolone antibiotic; penam	P	A
acrS	antibiotic efflux; cephalosporin; cephamycin; fluoroquinolone antibiotic; glycylcycline; penam; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic; triclosan	P	A
ADC-78	antibiotic inactivation; cephalosporin	P	A
adeB	antibiotic efflux; glycylcycline; tetracycline antibiotic	P	A

adeF	antibiotic efflux; fluoroquinolone antibiotic; tetracycline antibiotic	A	A
adeI	antibiotic efflux; carbapenem; cephalosporin; diaminopyrimidine antibiotic; fluoroquinolone antibiotic; lincosamide antibiotic; macrolide antibiotic; penem; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic	P	A
adeJ	antibiotic efflux; carbapenem; cephalosporin; diaminopyrimidine antibiotic; fluoroquinolone antibiotic; lincosamide antibiotic; macrolide antibiotic; penem; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic	P	A
adeK	antibiotic efflux; carbapenem; cephalosporin; diaminopyrimidine antibiotic; fluoroquinolone antibiotic; lincosamide antibiotic; macrolide antibiotic; penem; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic	P	A
ANT(4')-Ib	aminoglycoside antibiotic; antibiotic inactivation	P	P
APH(3'')-Ib	aminoglycoside antibiotic; antibiotic inactivation	P	P
APH(3')-Ia	aminoglycoside antibiotic; antibiotic inactivation	P	P
APH(3')-IIa	aminoglycoside antibiotic; antibiotic inactivation	P	P
APH(3')-IIIa	aminoglycoside antibiotic; antibiotic inactivation	P	P

APH(6)-Id	aminoglycoside antibiotic; antibiotic inactivation	P	A
apmA	aminoglycoside antibiotic; antibiotic inactivation	P	A
arnA	antibiotic target alteration; peptide antibiotic	P	A
bacA	antibiotic target alteration; peptide antibiotic	P	A
catB2	antibiotic inactivation; phenicol antibiotic	P	A
cfrA	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; oxazolidinone antibiotic; phenicol antibiotic; pleuromutilin antibiotic; streptogramin antibiotic	P	P
CRP	antibiotic efflux; fluoroquinolone antibiotic; macrolide antibiotic; penam	A	P
dfrC	antibiotic target replacement; diaminopyrimidine antibiotic	P	P
dfrG	antibiotic target replacement; diaminopyrimidine antibiotic	P	P
dfrK	antibiotic target replacement; diaminopyrimidine antibiotic	P	A
efmA	antibiotic efflux; fluoroquinolone antibiotic; macrolide antibiotic	P	A
emrB	antibiotic efflux; fluoroquinolone antibiotic	A	P
emrY	antibiotic efflux; tetracycline antibiotic	A	P
Enterococcus faecium chloramphenicol acetyltransferase	antibiotic inactivation; phenicol antibiotic	P	P

erm(44)	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic	A	P
ErmA	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic	P	P
ErmB	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic	P	A
ErmC	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic	P	P
ErmT	antibiotic target alteration; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic	P	A
Escherichia coli *acrA*	antibiotic efflux; cephalosporin; fluoroquinolone antibiotic; glycylcycline; penam; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic; triclosan	P	A
Escherichia coli *ampC*	antibiotic inactivation; cephalosporin; penam	P	P
Escherichia coli mdfA	antibiotic efflux; benzalkonium chloride; rhodamine; tetracycline antibiotic	P	A
evgS	antibiotic efflux; fluoroquinolone antibiotic; macrolide antibiotic; penam; tetracycline antibiotic	P	P
fexA	antibiotic efflux; phenicol antibiotic	P	P
FosB3	antibiotic inactivation; fosfomycin	P	P
fusB	antibiotic inactivation; fusidic acid	P	P

Lactobacillus reuteri cat-TC	antibiotic inactivation; phenicol antibiotic	P	A
lmrP	antibiotic efflux; lincosamide antibiotic; macrolide antibiotic; streptogramin antibiotic; tetracycline antibiotic	P	A
lnuA	antibiotic inactivation; lincosamide antibiotic	P	P
lsaB	antibiotic target protection; lincosamide antibiotic; pleuromutilin antibiotic; streptogramin antibiotic	P	A
mdtB	aminocoumarin antibiotic; antibiotic efflux	A	P
mdtC	aminocoumarin antibiotic; antibiotic efflux	P	P
mdtN	acridine dye; antibiotic efflux; nucleoside antibiotic	P	P
mdtC	acridine dye; antibiotic efflux; nucleoside antibiotic	P	P
mdtO	acridine dye; antibiotic efflux; nucleoside antibiotic	A	P
mecA	antibiotic target replacement; carbapenem; cephalosporin; cephamycin; monobactam; penam	P	P
mecC	antibiotic target replacement; carbapenem; cephalosporin; cephamycin; monobactam; penam	P	A
mepA	antibiotic efflux; glycylcycline; tetracycline antibiotic	P	P
mepR	antibiotic efflux; glycylcycline; tetracycline antibiotic	P	P

MexF	antibiotic efflux; diaminopyrimidine antibiotic; fluoroquinolone antibiotic; phenicol antibiotic	P	A
mgrA	acridine dye; antibiotic efflux; cephalosporin; fluoroquinolone antibiotic; penam; peptide antibiotic; tetracycline antibiotic	P	P
mphC	antibiotic inactivation; macrolide antibiotic	P	P
msbA	antibiotic efflux; nitroimidazole antibiotic	P	P
msrA	antibiotic target protection; macrolide antibiotic; streptogramin antibiotic	P	P
msrE	antibiotic target protection; macrolide antibiotic; streptogramin antibiotic	P	A
mupA	antibiotic target alteration; mupirocin	P	P
norA	acridine dye; antibiotic efflux; fluoroquinolone antibiotic	A	P
oqxB	antibiotic efflux; diaminopyrimidine antibiotic; fluoroquinolone antibiotic; glycylcycline; nitrofuran antibiotic; tetracycline antibiotic	P	A
OXA-214	antibiotic inactivation; cephalosporin; penam	A	P
OXA-72	antibiotic inactivation; cephalosporin; penam	P	A
blaZ	antibiotic inactivation; penam	P	P
qacA	antibiotic efflux; fluoroquinolone antibiotic	P	P
qacB	antibiotic efflux; fluoroquinolone antibiotic	P	P
QnrB10	antibiotic target protection; fluoroquinolone antibiotic	P	A

salA	antibiotic target protection; lincosamide antibiotic; pleuromutilin antibiotic; streptogramin antibiotic	P	A
SAT-4	antibiotic inactivation; nucleoside antibiotic	P	P
SHV-104	antibiotic inactivation; carbapenem; cephalosporin; penam	P	A
spd	aminoglycoside antibiotic; antibiotic inactivation	P	A
Staphylococcus intermedius chloramphenicol acetyltransferase	antibiotic inactivation; phenicol antibiotic	P	A
Streptococcus suis chloramphenicol acetyltransferase	antibiotic inactivation; phenicol antibiotic	P	A
TEM-116	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	P
TEM-149	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-162	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-171	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-193	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-201	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	P
TEM-207	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-220	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
TEM-40	antibiotic inactivation; cephalosporin; monobactam; penam; penem	P	A
tet(38)	antibiotic efflux; tetracycline antibiotic	P	A

tet(C)	antibiotic efflux; tetracycline antibiotic	P	P
tet(K)	antibiotic efflux; tetracycline antibiotic	P	P
tet(L)	antibiotic efflux; tetracycline antibiotic	P	A
tet(W/N/W)	antibiotic target protection; tetracycline antibiotic	P	P
tetM	antibiotic target protection; tetracycline antibiotic	P	A
tetR	antibiotic efflux; antibiotic target alteration; glycylcycline; tetracycline antibiotic	P	A
tetS	antibiotic target protection; tetracycline antibiotic	P	A
tetT	antibiotic target protection; tetracycline antibiotic	P	A
tolC	aminocoumarin antibiotic; antibiotic efflux; cephalosporin; cephamycin; fluoroquinolone antibiotic; glycylcycline; macrolide antibiotic; penam; phenicol antibiotic; rifamycin antibiotic; tetracycline antibiotic; triclosan	A	P
vanA	antibiotic target alteration; glycopeptide antibiotic	P	A
vanHA	antibiotic target alteration; glycopeptide antibiotic	P	A
vanRA	antibiotic target alteration; glycopeptide antibiotic	P	A
vanSA	antibiotic target alteration; glycopeptide antibiotic	P	A
vanXA	antibiotic target alteration; glycopeptide antibiotic	P	A

vanYA	antibiotic target alteration; glycopeptide antibiotic	P	A
vanZA	antibiotic target alteration; glycopeptide antibiotic	A	A
vatB	antibiotic inactivation; streptogramin antibiotic	A	P
vgaA	antibiotic target protection; pleuromutilin antibiotic; streptogramin antibiotic	P	P
vgaALC	antibiotic target protection; pleuromutilin antibiotic; streptogramin antibiotic	A	P
vgaB	antibiotic target protection; pleuromutilin antibiotic; streptogramin antibiotic	P	P
vgaE	antibiotic target protection; pleuromutilin antibiotic; streptogramin antibiotic	p	A
yojI	antibiotic efflux; peptide antibiotic	P	A

Table 1.5: List of antibiotic resistance genes and function that have been identified in *S. aureus* and *S. epidermidis* resistome. P=present, A=absent . This table only indicates if the antibiotic resistant genes were detected in staphylococci genome and plasmids from WGS data from the NCBI. This table does not indicate if the genes will encode antibiotic resistance for staphylococci. Data from **https://card.mcmaster.ca/download updated 20 March 2019**.

1.7.1 Antibiotic resistance mechanisms

Bacteria have developed sophisticated mechanisms to avoid being killed by antibiotics. These mechanisms are chemical alterations or destruction of the antibiotic using enzymes; decreased antibiotic penetration and efflux and the change of target site (Figure 1.5) (Munita and Arias, 2016). These mechanisms have evolved over millions of years, of which resistance to a single antimicrobial class may be caused by multiple changes in biochemical pathways. For instance, fluoroquinolone resistance can accrue by either mutation in the genes that fluoroquinolone targets; over-expression of the efflux pump and protection of fluoroquinolone targets site with a protein called Qnr (Munita and Arias, 2016).

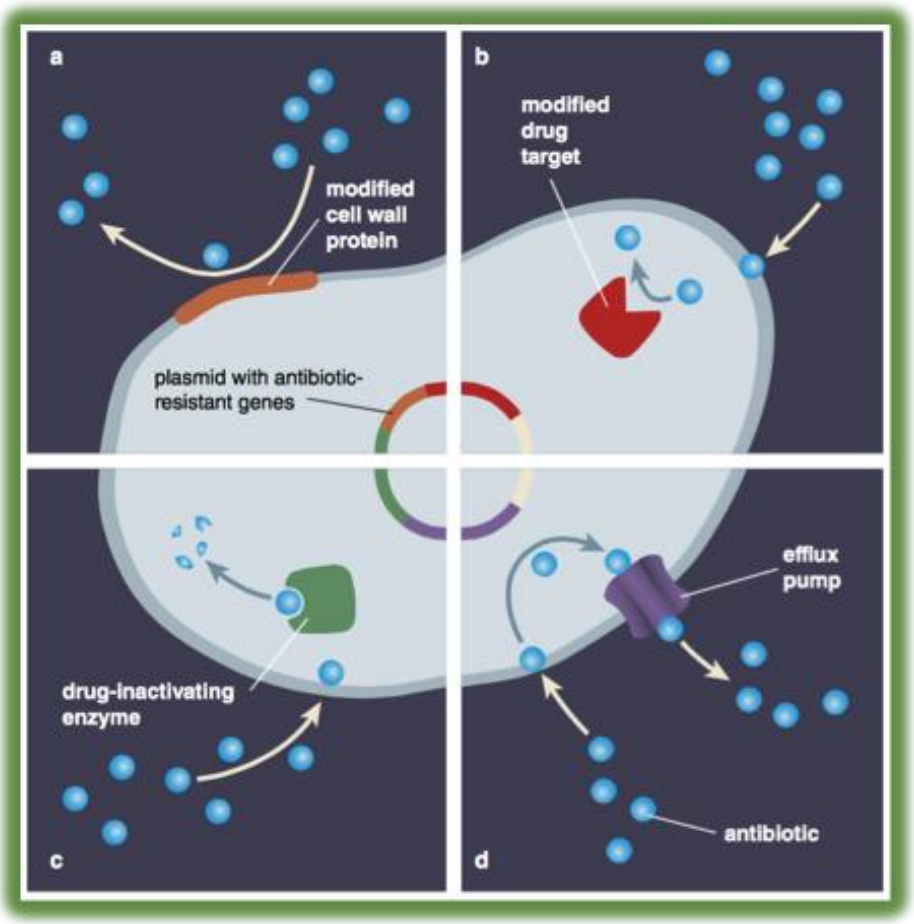


Figure 1.5: Mechanisms that have evolved in bacterial antimicrobial resistance. Figure from Dantas and Sommer, 2014

1.7.1.1 Penicillin resistance

A common antibiotic-resistant mechanism feature found in staphylococci is penicillinase enzyme. This enzyme targets the beta-lactam rings found in penicillin, breaking down the molecular structure and therefore stopping the antibiotic binding to its target site. In staphylococci the *blaZ* gene is responsible for producing penicillinase enzyme. This gene is commonly identified in staphylococci from hospitals; community; animal and the environment. *BlaZ* has been identified to be encoded on the chromosome or plasmid in staphylococci. Penicillinase was identified as part of first epidemiology antibiotic resistance wave in hospital in *S. aureus* in the 1940s where it is believed to be encoded on a plasmid which can be horizontally transferred to other *S. aureus* or staphylococci sensitive penicillin (Chambers and Deleo, 2009).

1.7.1.2 Broad spectrum resistance to beta-lactam antibiotics

The second epidemic wave of antibiotic resistance in *S. aureus* came shortly after the mass use of methicillin in hospitals in the 1950s (Chambers and Deleo, 2009). Methicillin was a semisynthetic derivative of penicillin which was not affected by beta-lactamase. Instead *S. aureus* gains resistance to this antibiotic by replacing the drug target PBP2 protein to the PBP2A protein. This new protein lowers the binding ability of drugs that have the β-lactam ring and give resistance to multiple drug classes, including carbapenem, cephalosporin, penam, cephamycin and monobactam (Chambers and Deleo, 2009).

1.7.1.3 Staphylococcal Chromosomal Cassette *mec*

The *mecA* gene in staphylococci can be located on a mobile genetic element known as Staphylococcal Chromosomal Cassette *mec* (SCC*mec*) (Oliveira et al., 2002). This mobile element is about 60kb in length and inserted itself next to the *orfX* gene (Boundy

et al., 2013; Oliveira et al., 2002). The SCC*mec* consist of two regions known as the *mec* complex and the chromosome cassette recombinase (*ccr*) complex. These regions can be genetically variable and normally used in classifying between different SCC*mec* types. To date, there are 11 (I-XI) different SCC*mec* types with further subtypes which are organised into a hierarchical system (Elements (IWG-SCC), 2009) (Figure 1.6). The different *mec* complexes are assigned A, B, C1, C2, D, E and the *ccr* complexes as A1/B1, A2B2, A3/B3, A4/B4, C1, A5/B3, A1/B6, A1 (Elements (IWG-SCC), 2009; Li et al., 2011; Shore et al., 2011). These *SCCmec* types normally contain one *mec* complex and *ccr* complex though there are reports of unclassified SCC*mec* that have multiple complexes isolated from clinical and community sites (Chen et al., 2017). The third wave of antibiotic resistance epidemic in *S. aureus* in the 1970s was associated with SCC*mec* type II and SCC*mec* type III, whereas wave four happened in the mid-1990s and was marked by the spread of MRSA in the community with a smaller mobile element known as the SCC*mec* type IV (Chambers and Deleo, 2009).

1.7.1.3.1 *mec* complex

The *mec* complex normally carries the *mecA* gene along with its two regulatory genes *mecR* and *mecI* (Petinaki et al., 2001). These genes can be flanked by insertion sequences (IS) (Noto et al., 2008). *MecR* and *mecI* are not always present in the *mec* complex as it has been reported that only 60–95% of MRSA isolates have them (Petinaki et al., 2001). It has also been noted that the regulatory gene and IS have been identified as being truncated in some *mec* complexes (Shore et al., 2005).

1.7.1.3.2 *ccr* complex

The *ccr* genes are located in the *ccr* complex that is responsible for making the SCC*mec* element mobile (Ito et al., 2001). To date, three distinct *ccr* gene alleles have been discovered named *ccrA*, *ccrB* and *ccrC* of which *ccrA* and *ccrB* can be classified into four different allotypes (Elements (IWG-SCC), 2009). *Ccr* is grouped in the same allotypes

sharing 85% or more of the same nucleotides, whereas *ccr* in different allotypes normally share 60-82% of the same nucleotide (Elements (IWG-SCC), 2009).

1.7.1.3.3 J-region

Areas between mec and *ccr* complex are known as the joint region (j-region) (Elements (IWG-SCC), 2009). These regions are used for subtyping different SCC*mec* classes (Elements (IWG-SCC), 2009). In these regions, it has been reported to have known characteristic genes, pseudogenes, or noncoding regions as well as other mobile genetic elements (Ito et al., 2007). These areas have been shown to sometimes carry additional resistance genes including aminoglycosides, macrolides, fusidic acid and heavy metal ions (Lin et al., 2014; Monecke et al., 2011).

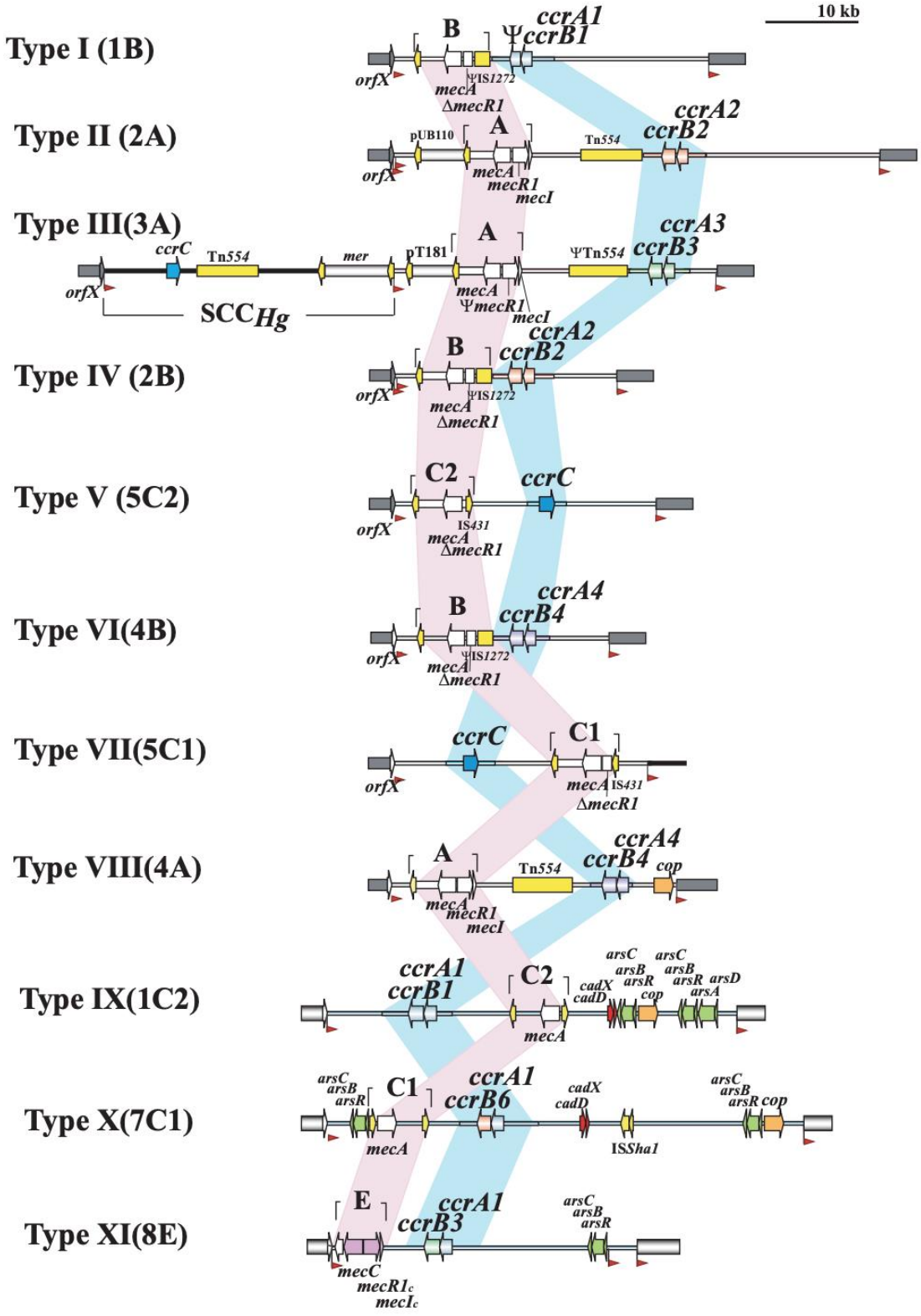


Figure 1. 6: SCC*mec* classification. The structural organization of SCC*mec* elements based on their *mec* complex (purple) and *ccr* complex (blue). In between these complexes are the J-region. Figure from Kaya et al., 2018.

1.7.1.3.4 *mecC* gene

The *mecA* was long thought to be one conserved gene, however in 2011 a new emergent gene known as *mecC* was discovered which has 69% nucleotide and 63% amino acid homology to *mecA* (Kerschner et al., 2014). Originally this was a problem as this gene caused *S. aureus* infection to be misdiagnosed as methicillin-sensitive due to the fact they are more susceptibility to oxacillin and the PCR diagnosis for MRSA would not work due to new gene low homology toward *mecA* (Paterson et al., 2014). Interestingly, *mecC* isolates had increase susceptibility to cefoxitin.

Little is known about the origin of the SCC*mec* element although there is some evidence that it might have originated in *S. fleurettii* (Tsubakishita et al., 2010). Interestingly they found that *S. fleurettii* had the *mecA* gene, but it did not contain the rest of the SCC*mec* element (Tsubakishita et al., 2010). This species is typically found as a commensal strain in animals, which has shown no clinical infection in humans (Tsubakishita et al., 2010). *S. fleurettii* cannot uptake the SCC*mec* element due to already having resistance to β-lactam ring antibiotics. It is most likely that this species evolved resistance from the environment from an antibiotic produced by a fungus or by the medication/growth supplements given to livestock (Tsubakishita et al., 2010).

1.7.1.4 Aminoglycoside-modifying enzymes

Aminoglycoside-modifying enzymes are a large family of enzymes that have been previously identified in modifying the molecular structure of aminoglycoside antibiotics. This family of enzymes can be further split into three subclasses based on the type of modification they cause. These classes are aminoglycoside *N*-acetyltransferases (AACs); aminoglycoside *O*-nucleotidyltransferases (ANTs), and aminoglycoside *O*-phosphotransferases (APHs). There are members of these families that are bifunctional *AAC(6')-Ie/APH(2")-Ia* from *S. aureus* (Garneau-Tsodikova and Labby, 2016).

1.7.1.4 Vancomycin resistance

Vancomycin resistance in *Staphylococcus* is caused by the decreased of antibiotic penetration. Reports in *S. aureus* show that resistance to vancomycin is due to increase peptidoglycan synthesis making the cell shape irregular, increase in cell wall thickness, a decrease in cross-linkage of peptidoglycan strands which revealed more D-Ala-D-Ala residues (Hanaki et al., 1998). The increase in D-Ala-D-Ala residues bind and trap vancomycin, preventing it from reaching its target in the cytoplasmic membrane (Hanaki et al., 1998). The second form of resistance has also been described is the *vanA* operon, which is conjugal transfer on a plasmid from vancomycin-resistance *Enterococcus faecalis*. This alters D-Ala-D-Lac residues which reduces the binding affinity with vancomycin (González-Zorn and Courvalin, 2003).

1.7.1.5 Efflux pumps

Efflux pumps are an assembly of proteins involved in the removal of a single or multiple toxic molecules out of bacterial cells (Webber and Piddock, 2003). They are grouped into 5 structural families; the resistance-nodulation-division (RND), the small multidrug resistance (SMR), the multi antimicrobial extrusion (MATE), the major facilitator superfamily (MFS), and the ATP-binding cassette (ABC) superfamilies (Hernando-Amado et al., 2016.) In Gram-positives bacteria, some efflux pumps can work independently of any other protein whereas, in Gram-negative, they form tripartite complexes capable to traverse both bacterial membranes (Alcalde-Rico et al., 2016). Some of these efflux pumps can remove multiple types of antibiotics and are known as a multidrug efflux pump. A common efflux transporter in staphylococci that transports a single type of molecule is *tet(K)* which can actively transport tetracycline whereas *norA* is multidrug efflux transporter found in *S. epidermidis* that targets fluoroquinolone and acridine dye (Costa et al., 2019, Yamaguchi et al. 1995).

1.8 Aims

This thesis aims were to recover and characterise environmental MDR staphylococci from high-frequency hand touched surfaces from general public settings and public areas in hospitals from East and West London and compare these isolates to other known clinical, animal, plant and environmental associated lineages by performing comparative whole genome sequencing analysis. These results will aid in the understanding of the evolution of antibiotic resistance and virulence in the environmental MDR staphylococci found. The project aims are to determine:

1. the phenotypic and genetic background of environmental multidrug-resistant staphylococcal isolates,

2. the factors distinguishing them from one another and any mutations which occur allowing environmental isolates to acquire the resistance and virulence genes,

3. whole genome sequencing of environmental multidrug-resistant isolates which will be compared with well-characterised reference strains.

These objectives will use a "One Health" approach which will allow us to determine if the genetic lineages of MDR staphylococci from public settings are genetically related to isolates from clinical, non-clinical, animal or environmental lineages. These results will help to determine/reveal if isolates from general public settings could pose a public health risk. In this study, two different geographical areas with similar features (shopping centre, train station and hospitals) were compared to identify differences in the abundance of MDR staphylococci in these areas.

This thesis is comprised of 7 chapters. Chapter 1 is comprised of the introduction; which discusses previous research in relation to the research project; Chapter 2 is the material and methods used in performing experimental procedures. Chapter 3, 4, 5 and 6 are on

the results and how they relate to other studies and the implications of these finding and Chapter 7 has the overall conclusion from the thesis and ideas for future works.

Chapter 2: Material and methods

2.1 Sample collection

Staphylococcal isolates were recovered from high-frequency hand touched surfaces of inanimate objects (door handles, stair handrails, toilet flushers, toilet seats, taps, lift buttons, chair armrests) from four locations in general public settings, two locations from East London and two locations from West London between November 2016 to September 2017 (figure 2.1). Public settings included shopping centres (concourses, escalators lifts, public washrooms) and train stations (entry gates, public washrooms, escalators). Isolates were also recovered from a hospital setting where the general public had easy access, without being a patient or visiting a patient (reception area, public washrooms, corridors, lifts). These sampling areas from East and West London were chosen as both have a large shopping centre, train station and hospital in close proximity to each other. From each location, 50 sites were randomly sampled using COPAN dry swabs (Copan Diagnostics Inc., USA). In total 600 isolates were recovered of which 224 were from East London and 376 from West London of which 182 of the isolates were from the community area and 418 from hospital areas; 97 from East London community area and 85 from West London community and 224 from East London hospital and 376 from West London hospital.

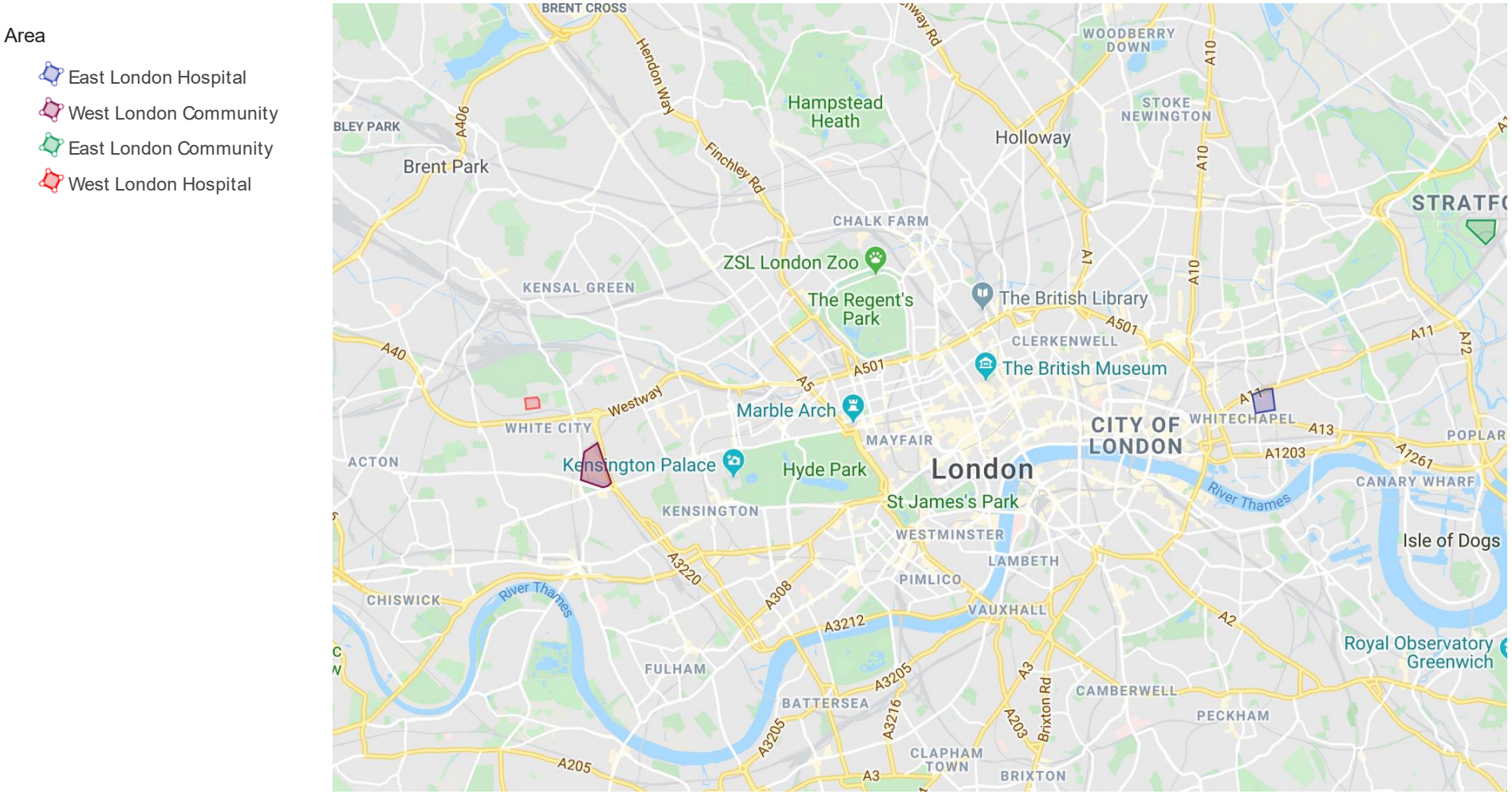


Figure 2.1: A map of sampling sites in East and West London.

2.2 Isolation of staphylococci

All samples were directly inoculated onto mannitol salt agar (MSA, Oxoid Basingstoke, UK) within 1-3 hours of recovery and incubated aerobically for 24-72 hours at 37°C. MSA agar was prepared from powder and mixed into distilled water before being sterilised in an autoclave at 121ºC for 15 minutes. Molten agar was poured into a 15mm x 100mm agar plate and was left to solidify. CoPS growth on MSA would produce yellow colonies with yellow halos where CoNS growth on MSA would produce pink colonies and no change to media colour (Figure 2.2). To prevent bias, up to 10 colonies from each plate were picked each having different colony morphology or if there are less than 10 different colony morphologies an equal amount of different colony morphologies was selected. These isolates were screened for potential staphylococci characteristics, including performing catalase and coagulase tests. Prolex™ staph latex kits (ProLab Diagnostics, Neston, UK) was used to distinguish *S. aureus* and coagulase-negative staphylococci.

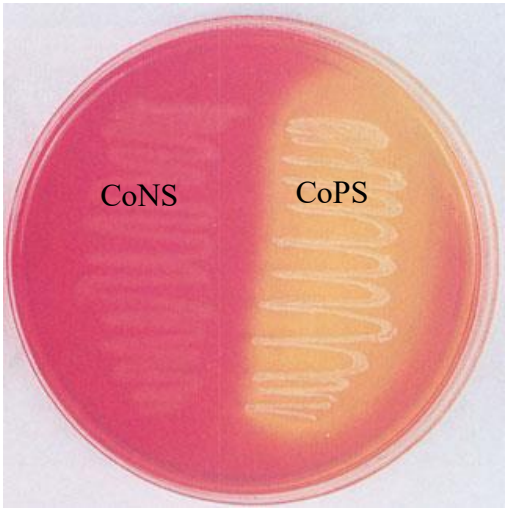
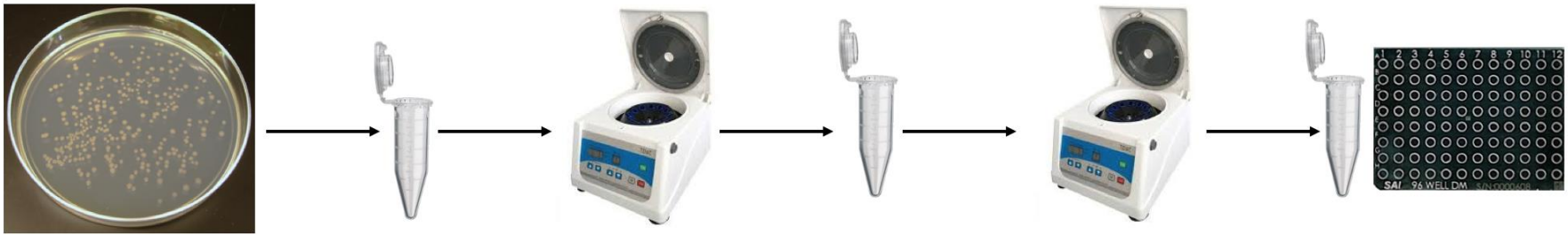


Figure 2.2: Mannitol salt agar and colony morphology between CoPS and CoNS

2.3 Identification of staphylococci recovered from high-frequency hand touch areas

Potential staphylococcal isolates were initially identified by conventional methods, including Gram staining and catalase testing. All the isolates were identified at species level using Matrix-assisted laser desorption ionisation-time of flight mass spectroscopy (MALDI-TOF-MS, Microflex LT, Bruker Daltonics, Coventry, UK) in a positive linear mode (2000–20,000 m/z range) (figure 2.3). Samples were prepared by growing bacterial cells on nutrient agar plates (prepared in the same manner as the MSA agar) at 37°C. A few colonies were picked from the agar plate and resuspended in a microcentrifuge tube containing 300 µl of sterile distilled water, which was then mixed with 900 µl of absolute ethanol. The microcentrifuge tube was then centrifuged for 2 minutes at max speed and the supernatant was removed by pipetting. The pellet was then resuspended in 50 µl of 70% formic acid (Sigma-Aldrich, UK) before 50 µl of acetonitrile (Sigma-Aldrich, UK) was mixed into the solution. The solution was then centrifuged at max speed for 2 minutes and the 1 µl of supernatant was placed onto an MSP 96 target polished steel plate (Bruker Daltonics, Coventry, UK) and left to air-dry. 1 µl α-cyano-4-hydroxycinnamic acid (HCCA) (Bruker Daltonics, Coventry, UK) used for matrix solution was overlaid. MALDI-TOF Biotyper 3.0 software (Bruker Daltonics, Coventry, UK) was used to analyse the spectra and to identify the bacterial species. Bacterial test standard *Escherichia coli* DH5α (Bruker Daltonics, Coventry, UK) was used for calibration and as a standard for quality control. Isolates that had higher confidence of identification had a score > 2.0 and low confidence of 1.7-1.90.



1. A few colonies were picked from the agar plate.

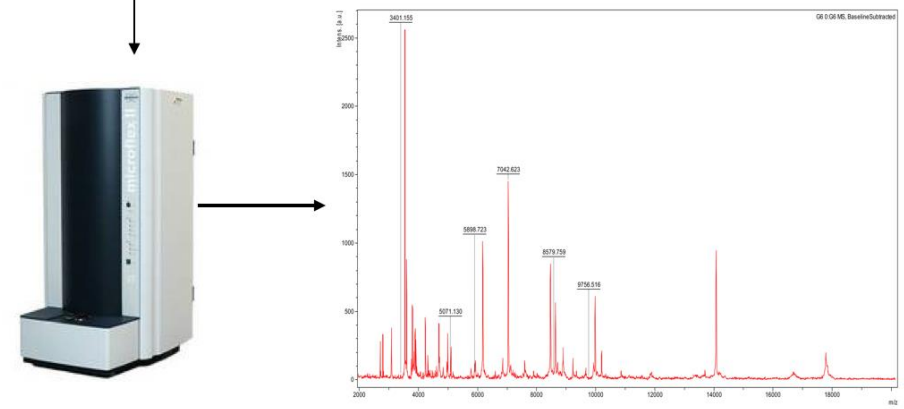
2. Colonies resuspended in a microcentrifuge tube containing 300 µl of sterile distilled water, which was then mixed with 900 µl of absolute ethanol

3. Microcentrifuge tube was then centrifuged for 2 minutes at max speed

4. The supernatant was removed by pipetting. The pelleted was then resuspended in 50 µl of 70% formic acid before 50 µl of acetonitrile was mixed into the solution.

5. Microcentrifuge tube was then centrifuged for 2 minutes at max speed.

6. 1 µl of supernatant was placed onto a steel target plate and left to air-dry. Dried samples was overlaid with 1 µl HCCA.



7. Target plate analysed in MALDI-TOF MS.

8. Spectra was compared with a database of known bacterial species mass ions to identify bacterial species

Figure 2.3: Schematic on how staphylococcal species were identified by MALDI-TOF

2.4 Comparative identification analysis of staphylococcal species using the Bruker Autoflex and ASTA Tinkerbell MALDI-TOF MS

The accuracy and reproducibility of MALDI-TOF MS were tested by comparing identification data of the Brucker Autoflex (Biotyper 3.0 software) with the ASTA Tinkerbell (micro ID software) (Figure 2.4). A single colony was picked from a plate with a sterile toothpick and placed onto a well of 384 circles of the µFocus MALDI Plate 2000 µm (ASTA, Manchester, UK). 1 µl of 70% formic acid was added on top of the colony on the target pate. The plate was left to air dry before the HCCA used for matrix solution was overlaid. The same target plate with the same spotted isolates was analysed by both the Bruker's Autoflex and the ASTA's Tinkerbell software. Bacterial test standard *Escherichia coli* DH5α (Bruker Daltonics, Coventry, UK) was used for calibration and as a standard for quality control. Both instrument parameters were operated at linear mode and both measured a mass range –m/z 200-20,000 Da. The Autoflex instrument ion extraction voltage is 19.5kV; fires 1,000 laser shots per spot and frequency of 200Hz whereas the Tinkerbell instrument sample voltage is 18kV; fires 1,200 laser shot per sample and has a delay time of 1.1 µs. For the Tinkerbell micro ID software, a score of >140 indicated high confidence identification and 110 to 139 low confidence of identification. Isolates were tested in triplicate. Spectra from both machines were compared using Matlab software to identify similarities in mass (Natick, Massachusetts, USA).

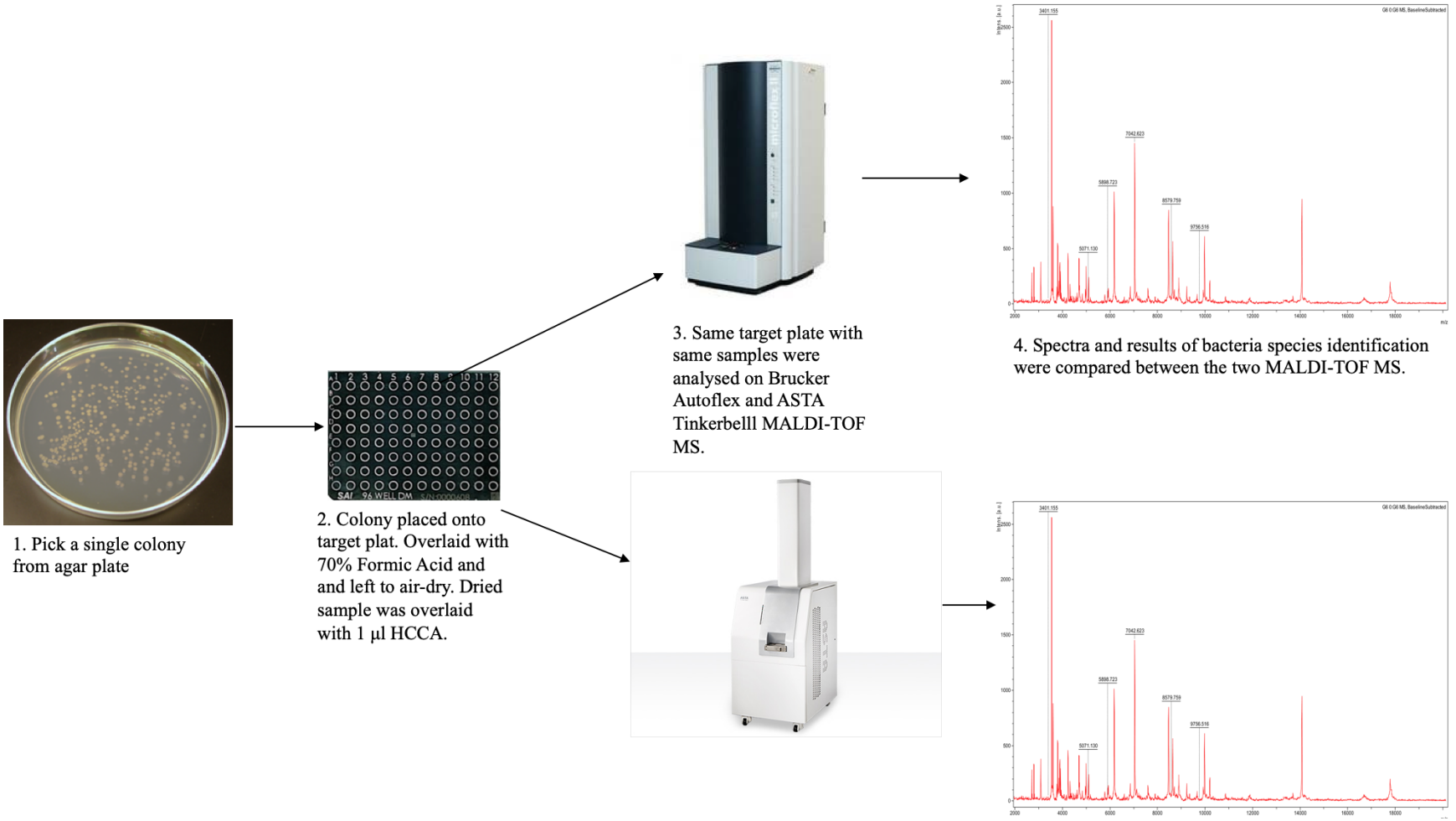


Figure 2.4: Schematic on the comparative analysis of staphylococcal species identification using the Bruker Autoflex and ASTA Tinkerbell MALDI-TOF MS

2.5 Antimicrobial susceptibility testing

All isolates were tested for their susceptibility against a panel of 11 antibiotics by using a disc diffusion method (Andrews and Howe, 2011). The antibiotics tested were the following: oxacillin (1µg), gentamicin (10 µg), mupirocin (20 µg), amoxicillin (10 µg), erythromycin (15 µg), tetracycline (10 µg), cefoxitin (30 µg), cefepime (30 µg), fusidic acid (10 µg), penicillin (1 unit) and chloramphenicol (30 µg) (Mast Group, Merseyside, UK). Antibiotic profiles of each isolate were determined according to the recommendation of the (CLSI) and British Society for Antimicrobial Chemotherapy (BSAC) (Andrews and Howe, 2011; CSLI, 2017). In addition, the minimum inhibitory concentrations (MIC) for oxacillin and cefoxitin were determined using E-tests (Biomerieux, Basingstoke) (Andrews and Howe, 2011; CSLI, 2017). Bacteria culture from MSA plates were streaked onto nutrient agar plates and grown aerobically for 16 to 24 hours at 37 ^{O}C. Samples from the agar plates were then suspended into 500 µl of nutrient broth (sterilised the same as MSA) and the turbidity was adjusted to the 0.5 McFarland standard. 100 µl of the adjusted inoculum was spread onto 15 ml Muller Hinton agar plates (prepared similarly to MSA) and then left to dry before aseptically placing the antibiotics discs or E-test strips onto agar plates. The inoculated Muller Hinton agar plates were then incubated for 24 hours at 37 ^{O}C. After incubation the zone of inhibition for disc diffusion and the minimum concentration that inhibits bacterial growth using the E-test was recorded (figure 2.5).

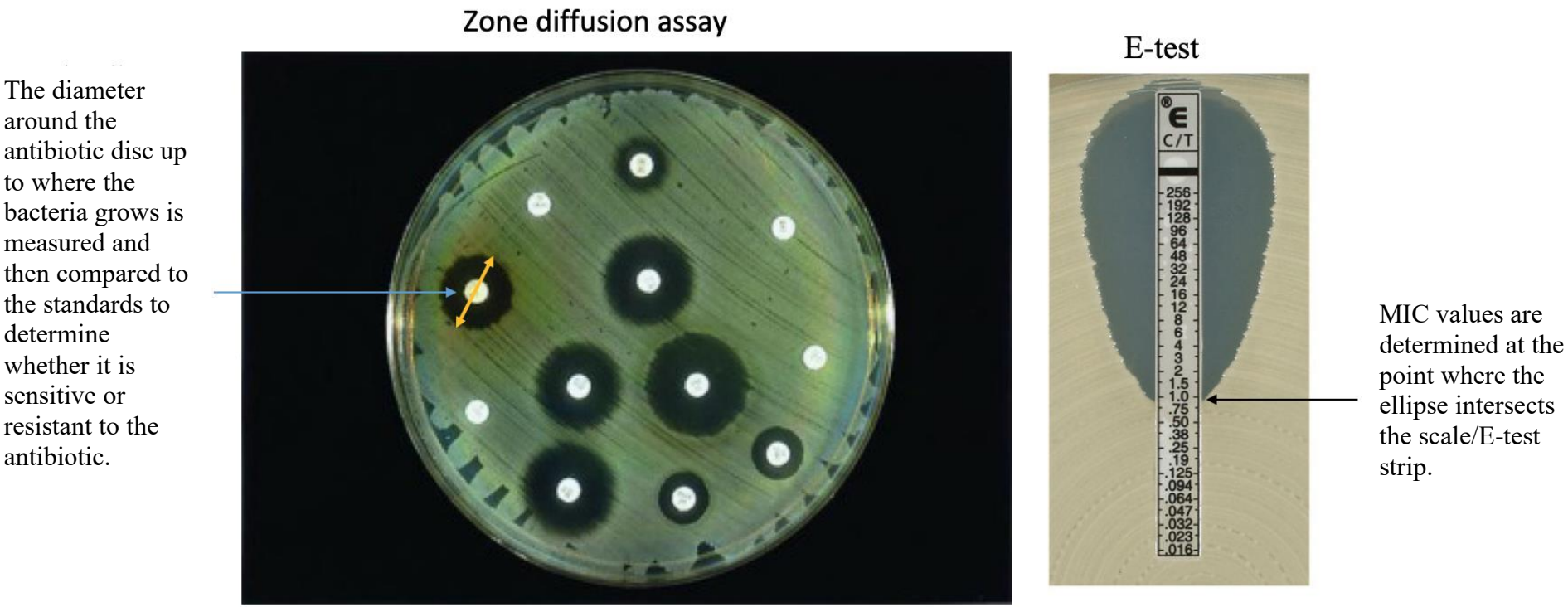


Figure 2.5: How to interpret inhibition zone of disc diffusion assay and E -test.

2.6 Detection of the *mecA* gene by PCR

The *mecA* gene was detected for all staphylococcal isolates using PCR. Freshly grown samples were suspended into 40 µl of sterile distilled water and boiled at 100 °C then cooled on ice for 5 minutes. The samples were then centrifuged at 13,000 × g for 1 minute and the supernatant was used for the PCR providing the DNA template. The PCR was performed using Met1 and Met2 primers (Eurofins, Germany) (Table 2.1). PCR reactions were performed in a 20 µl volume for each sample which consists of 10 µl of Phusion Master Mix;1 µl of met1, 1 µl of met2, 6 µl of sterile distilled water and 1 µl of isolates DNA template. The PCR condition can be found in Table 2.2. PCR product was loaded onto 1% agarose electrophoresis gel stained with SYBR safe (Thermofisher, UK) and ran at 120 V for 30 minutes. Electrophoresis gel was visualised using the ChemiDoc (Bio-Rad, UK) to determine the size of the PCR product against a 1 kb DNA ladder.

Prime	Sequence
Met1 Met2	5'-GGGATCATAGCGTCATTATTC-3' 5'-ACGATTGTGACACGATAGCC-3'

Table 2.1: Primers sequence for amplifying the *mecA* gene

Step	Temperature (°C)	Time	Cycles
Initial denaturing	93	5 minutes	1
Denaturing	93	30 seconds	35
Annealing	52	30 seconds	
Extension	72	1 minute	
Final Extension	72	10 min	1

Table 2.2: PCR conditions for *mecA* gene

2.6 WGS and bioinformatic analysis

For a flow chart of all software used for genome assemble and bioinformatic analysis see figure 2.6.

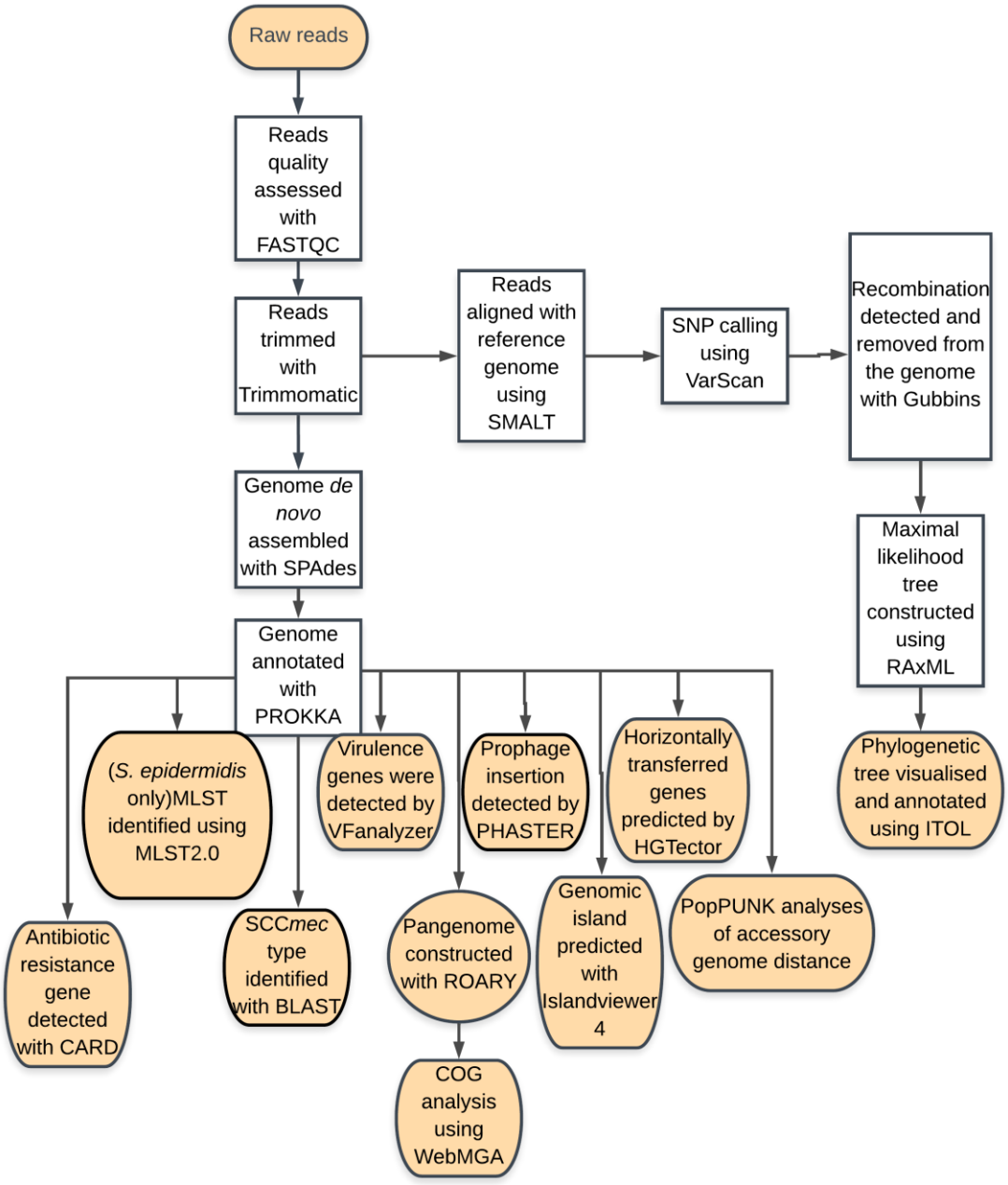


Figure 2.6: Flow chart of bioinformatic software used in genome assembly and bioinformatic anaylsis.

2.6.1 Genome sequencing

49 staphylococci *mecA* positive (*mecA*$^{+}$) isolates were whole genome sequenced using Illumina HiSeq platforms. 13 out of 49 isolates were whole genome sequenced by MicrobesNG (Birmingham, UK) and the remaining isolates were sequenced at Fudan University, Shanghai, China.

Genomic DNA was extracted using TIANamp Bacteria DNA kit (Tiangen, China) and paired-end sequencing libraries were constructed using Nextera XT DNA Sample Preparation kits or TruSeq DNA HT Sample Prep Kit (Illumina, USA) following the manufacturer's instruction. The sequence coverage for each read was set for 30x.

2.6.2 Genome assembly

The raw reads quality was assessed using FASTQC software (Andrews, 2011). FASTQC measure per base sequence quality, per sequence quality scores, per base sequence content, per base GC content, per sequence GC content, per base N content, sequence length distribution, duplicate sequences, overrepresented sequences and overrepresented kmers (Andrews, 2011). From these measurements, the software can raise a warning if the reads are poor quality (Andrews, 2011). From the information given by FASTQC, the reads were trimmed using trimmomatic software (Version 0.35) set at Phred cutoff of Q20 to remove the miscalled bases from the end of the reads (Bolger et al., 2014).

The trimmed reads were *de novo* assembly by SPAdes 3.11, a fast and flexible software which uses k-mers and De Bruijn graph to assemble draft genomes into contigs without a reference genome (Bankevich et al., 2012). From the assembled genomes; contigs that were ≤ 500bp were removed as these contigs are missed assembled due to composition of only unpaired reads, wrong paired-end orientation and abundance of abnormal insert sizes between paired reads. QUAST was used to asses contig assembly by measuring the

number of contigs per genome, size of the largest contig, size of genome GC content, N50 and L50 (Gurevich et al., 2013).

2.6.3 Identifying antibiotic resistance and virulence genes from whole genome sequences

Antibiotic resistance genes were detected using the Comprehensive Antibiotic Resistance Database (CARD) server (June 2019) due to having a comprehensive up to date database of antibiotic resistance genes (Jia et al., 2017). The cutoff for antibiotic resistance genes identification was set at ≥95% similarity of DNA sequence. Virulence factors were detected by VFanalyzer server (June 2019) due to having the most comprehensive database of bacterial virulent factor genes and can avoid potential false positives due to paralog (Liu et al., 2019).

2.6.4 SCC*mec* screening

The diversity of SCC*mec* types were determined by searching against a database of known *SCCmec* molecular markers with NCBI BLAST with a cutoff e-value of 10^{-5} (Altschul et al., 1990; Monecke et al., 2016). NCBI BLAST is software that can find similarity between biological sequences accurately than other software (Altschul et al., 1990). The SCC*mec* marker database was from Monecke and co-workers study (Monecke et al., 2016).

2.6.5 Identification of mobile genetic elements

Plasmids were constructed from the trimmed read using PlasmidSPAdes (Antipov et al., 2016). This software can identify plasmids by the read coverage of contigs is higher or lower than the chromosome coverage (Antipov et al., 2016). The advantage of this

software in predicting plasmids is that it requires no prior knowledge of plasmid elements (Antipov et al., 2016).

Phage insertion and genomic island prediction were performed by PHASTER and Islandviewer4 (Arndt et al., 2016; Bertelli et al., 2017). PHASTER uses the most up to date database of phage protein and can accurately predict the completeness of prophages within genomes by using NCBI BLAST to identify what phage proteins are found within a region of the genome (Arndt et al., 2016). Islandviewer 4 use multiple software to predict genomic island against reference strains. The reference strains used in this study were *S. epidermidis* ATTC 11228; *S. haemolyticus* JCSC 1435 and *S. hominis* K1. The software used by Islandviewer 4 is SIGI-HMM (sequence composition prediction method using Hidden Markov Model and measures codon usage) and IslandPath-DIMOB (can detect abnormal sequence composition by dinucleotide bias composition or the presence of genes that functionally relate to mobile elements) (Bertelli et al., 2017).

Horizontally transferred genes and their predicted donor organism were predicted using HGTector pipeline (Zhu et al., 2014). This software uses the NCBI non-redundant protein database to BLAST sequences with cutoffs set at e-value=10^{-5} and percentage identity ≥ 30%, and query coverage ≥ 50%. "Self" group was determined by their species taxonomy ID (*S. epidermidis* ID 1282; *S. haemolyticus* ID 1283 and *S. hominis* ID 1290): “Close” group been *Staphylococcus* genus (taxonomy ID 1279) and a distal group containing all other species (Pruitt et al., 2007). The cutoff was determined for each species using a Gaussian kernel smoothing with Silverman's rule-of-thumb bandwidth selector as this works best for an approximately normal density distribution (Zhu et al., 2014). This software is better compared to other software’s due to being insensitive to stochastic events such as gene loss, rate variation and database error (Zhu et al., 2014).

2.6.6 MLST *S. epidermidis* isolates

S. epidermidis isolates MLST sequence types were assigned using MLST2.0 online service (Thomas et al., 2007). This server has the most up to date MLST allele sequence and profile data and sequences found in PubMLST.org (Thomas et al., 2007). This software was only used to type *S. epidermidis* isolates as there is no standardised MLST for other CoNS.

2.6.7 Core genome phylogenetic analysis

A core SNP Maximal likelihood tree was constructed using isolates recovered from different sources (Table 2.3). SMALT (version 0.5.8) (https://www.sanger.ac.uk/science/tools/smalt-0) was used to generate a hash index of the reference genomes and to map the reads to the reference genomes. A hash index was set with word length to 11 and sampling step size to 1. Reference genomes that were used to map each staphylococcal species were *S. epidermidis* ATTC 11228; *S. haemolyticus* JCSC 1435 and *S. hominis* K1. SMALT is preferred over other alignment software due to its flexibility and being more sensitive at detecting divergent hits (Caboche et al., 2014). SNP calling was done in parallel with all isolates of the same species using VarScan version 2.3.9. VarScan employs a robust heuristic/statistic approach to call variants compared with other software that mainly uses a statistical approach. (Koboldt et al., 2009). The VCF file was converted to multi-FASTA alignment file using a freely available python script vcf2phylip (https://github.com/edgardomortiz/vcf2phylip). Recombination was detected and removed from the genome using the software Gubbins set at default parameters (Croucher et al., 2015) Gubbins can accurate reconstructions under realistic models of short-term bacterial evolution and can rapidly process data (Croucher et al., 2015). A maximal likelihood tree was constructed using RAxML version 8 using the generalised time-reversible model (GTR) model with GAMMA method of

correction for site rate variation and 100 bootstrap replications (Stamatakis, 2014). RAxML is preferred over other software due to its handling of large datasets with its comparatively low memory consumption, advanced search algorithms and use of accelerated likelihood (Stamatakis, 2014). The phylogenetic tree was visualised and annotated using ITOL; an easy to use and free software on the internet (Letunic and Bork, 2016).

Isolate name	Species	Accession number	source	Reference
SE2.9	*S. epidermidis*	JRVN01000000	Rice seed	(Chaudhry and Patil, 2016)
SE4.7	*S. epidermidis*	JRVP01000000	Rice seed	(Chaudhry and Patil, 2016)
SE4.6	*S. epidermidis*	JRVO01000000	Rice seed	(Chaudhry and Patil, 2016)
SE4.8	*S. epidermidis*	JRVQ01000000	Rice seed	(Chaudhry and Patil, 2016)
UC7032	*S. epidermidis*	ARWU01000000	cured meat	(Gazzola et al., 2013)
CIM40	*S. epidermidis*	ATCW02000000	Mouse skin	(Wang et al., 2014)
APO27	*S. epidermidis*	ATCU02000000	Mouse skin	(Wang et al., 2014)
CIM28	*S. epidermidis*	ATDF02000000	Mouse skin	(Wang et al., 2014)
APO35	*S. epidermidis*	ATCV02000000	Mouse	(Wang et al., 2014)
NIHLM057	*S. epidermidis*	AKGO01000000	Human Occiput	(Conlan et al., 2012)
VCU128	*S. epidermidis*	AHLI01000000	Clinical airways	
NW32	*S. epidermidis*	LJIF01000000	Cow milk	
NIHLM015	*S. epidermidis*	AKGZ01000000	Human	(Conlan et al., 2012)
NIHLM037	*S. epidermidis*	AKGT01000000	Human	(Conlan et al., 2012)
14.1.R1	*S. epidermidis*	CP018842	Human skin	(Lassen et al., 2017)
MRSE 52-2	*S. epidermidis*	NTLC01000000	Human nasopharynx	(Magaña-Lizárraga et al., 2017)
NIHLM023	*S. epidermidis*	AKGU01000000	Human toe web	(Conlan et al., 2012)
S2 005 003 R3 50	*S. epidermidis*	QFPG01000000	hospital surfaces and sink	(Brooks et al., 2017)
M01	*S. epidermidis*	LYWE01000000	cowhouse	
ZSC	*S. epidermidis*	PHHR01000000	groundwater	
y24	*S. epidermidis*	NRSY01000000	bovine mastitis milk	
PR246B0	*S. epidermidis*	PCFD01000000	Pig rectum	

Isolate name	Species	Accession number	source	Reference
M25	*S. epidermidis*	LYWF01000000	cowhouse	
FDAARGOS-161	*S. epidermidis*	CP014132	Clinical Peripheral blood	
1457	*S. epidermidis*	CP020463	Central venous catheter	(Lassen et al., 2017)
AG42	*S. epidermidis*	JNLI01000000	Animal Sheep rumen	
Scl25	*S. epidermidis*	ATDC02000000	Mouse skin	(Wang et al., 2014)
CSF41498	*S. epidermidis*	CP030246	Clinical cerebrospinal fluid (meningitis)	(Galac et al., 2019)
SNUT	*S. epidermidis*	LQRB01000000	toluene treated bioreactor sludge	(Kim et al., 2016)
ATCC 12228	*S. epidermidis*	CP022247	Human skin and mucosal	(MacLea and Trachtenberg, 2017)
PM221	*S. epidermidis*	HG813242	Cow	(Savijoki et al., 2014)
SNUC 5038	*S. epidermidis*	PYYR01000000	Cow	(Naushad et al., 2016)
SNUC 75	*S. epidermidis*	PYZF01000000	Cow	(Naushad et al., 2016)
FDAARGOS_83	*S. epidermidis*	JTAY02000000	Clinical (urine)	
RP62A	*S. epidermidis*	CP000029	Clinical (intravascular catheter-associated sepsis)	(Gill et al., 2005)
SRR1182420	*S. epidermidis*	SRR1182420	Clinical blood	
ET-024	*S. epidermidis*	JGVL01000000	endotracheal tube biofilm of a mechanically ventilated patient	(Vandecandelaere et al., 2014)
SRR1182422	*S. epidermidis*	SRR1182422	Clinical blood	
SRR1182424	*S. epidermidis*	SRR1182424	Clinical blood	
SRR1182423	*S. epidermidis*	SRR1182423	Clinical blood	
SRR1182419	*S. epidermidis*	SRR1182419	Clinical blood	
SRR1182413	*S. epidermidis*	SRR1182413	Clinical blood	
VCU037	*S. epidermidis*	AFTY01000000	Clinical human airways	
M0881	*S. epidermidis*	AOAJ01000000	Human Skin	
VCU045	*S. epidermidis*	AFEI01000000	Clinical (human airways)	

Isolate name	Species	Accession number	source	Reference
FDAARGOS 153	*S. epidermidis*	CP014119	Clinical (peripheral blood)	
764 SEPI	*S. epidermidis*	JUTX01000000	Clinical	
M0026	*S. epidermidis*	JBVX01000000	Clinical (blood)	
NIH06004	*S. epidermidis*	AKHH01000000	Clinical(blood)	(Conlan et al., 2012)
NIH08001	*S. epidermidis*	AKHG01000000	Clinical (blood)	(Conlan et al., 2012)
SRR1182399	*S. epidermidis*	SRR1182399	Clinical blood	
SRR1182398	*S. epidermidis*	SRR1182398	Clinical blood	
SRR1182400	*S. epidermidis*	SRR1182400	Clinical blood	
SRR1182371	*S. epidermidis*	SRR1182371	Clinical blood	
SRR1182410	*S. epidermidis*	SRR1182410	Clinical blood	
SRR1182401	*S. epidermidis*	SRR1182401	Clinical blood	
SRR1182412	*S. epidermidis*	SRR1182412	Clinical blood	
SRR1182418	*S. epidermidis*	SRR1182418	Clinical blood	
SRR1182416	*S. epidermidis*	SRR1182416	Clinical blood	
BPH0662	*S. epidermidis*	LT571449	Clinical	
DAR1907	*S. epidermidis*	CP013943	Clinical (Blood)	
ENVH131	*S. epidermidis*	LYVR01000000	Hospital environment	
ENVH150	*S. epidermidis*	LYVW01000000	Hospital environment	
LRKNS114	*S. epidermidis*	LZEO01000000	Hospital environment	
LRKNS116	*S. epidermidis*	LZEQ01000000	Hospital environment	
LRKNS117	*S. epidermidis*	LZER01000000	Hospital environment	
SH06 17	*S. epidermidis*	PHKN01000000	Clinical blood	
SH03_17	*S. epidermidis*	PHKH01000000	Clinical (blood)	
SH07 17	*S. epidermidis*	PHKM01000000	Clinical (blood)	

Isolate name	Species	Accession number	source	Reference
760 SEPI	*S. epidermidis*	JUUB01000000	Clinical	
SNUC 3608	*S. epidermidis*	QXSP01000000	Cow	(Naushad et al., 2016)
SNUC 901.1	*S. epidermidis*	PYYQ01000000	Cow	(Naushad et al., 2016)
B45679-10	*S. epidermidis*	MVFV01000000	Clinical(blood)	
FDAARGOS 148	*S. haemolyticus*	LORN02000000	Clinical (blood)	
DNF00585	*S. haemolyticus*	JRNK01000000	Clinical (vagina)	
Z52	*S. haemolyticus*	PHHQ01000000	groundwater	
SW007	*S. haemolyticus*	MTIZ01000000	Dog	(Bean et al., 2017)
SNUC 128	*S. haemolyticus*	PZIV01000000	Cow	(Naushad et al., 2016)
SNUC 1317	*S. haemolyticus*	PZIP01000000	Cow	(Naushad et al., 2016)
SNUC 1450	*S. haemolyticus*	PZIL01000000	Cow	(Naushad et al., 2016)
C10F	*S. haemolyticus*	JQHA01000000	Clinical sputum	
SNUC 1408	*S. haemolyticus*	PZIM01000000	Cow	(Naushad et al., 2016)
2263-3461	*S. haemolyticus*	CUEO01000000	Clinical (teat)	(Cavanagh et al., 2014)
SNUC 4966	*S. haemolyticus*	PZIC01000000	Cow	(Naushad et al., 2016)
SNUC 1584	*S. haemolyticus*	PZIK01000000	Cow	(Naushad et al., 2016)
IPK TSA25	*S. haemolyticus*	NDWY01000000	surface area of a building with less than 200 occupants	
M-176	*S. haemolyticus*	CUEQ01000000	Clinical blood	(Cavanagh et al., 2014)
OG2	*S. haemolyticus*	NCXH01000000	Kefir seed	
RIT283	*S. haemolyticus*	JFOJ01000000	willow	
S167	*S. haemolyticus*	CP013911	Leaf vegetable	(Hong et al., 2016)
MTCC 3383	*S. haemolyticus*	LILF01000000	Human	

Isolate name	Species	Accession number	source	Reference
IIF2SW-P5	*S. haemolyticus*	MIZW01000000	waste and hygiene compartment of International Space Station	(Checinska Sielaff et al., 2016)
R1P1	*S. haemolyticus*	AJVA01000000	copper alloy coin	
SH1752	*S. haemolyticus*	LRHN01000000	Clinical infected eye	(Panda and Singh, 2016)
95671	*S. haemolyticus*	CUFA01000000	Central venous catheter	(Cavanagh et al., 2014)
SGAir0252	*S. haemolyticus*	CP025031	tropical air samples collected in Singapore	(Premkrishnan et al., 2018)
SHN36	*S. haemolyticus*	LRBN01000000	Healthy eye	(Panda and Singh, 2016)
BC05211	*S. haemolyticus*	MRUZ01000000	bovine milk	
8074328	*S. haemolyticus*	CUFG01000000	Clinical blood	(Cavanagh et al., 2014)
1HT3	*S. haemolyticus*	LAKG01000000	Clinical Colon	
115601	*S. haemolyticus*	CUHH01000000	Central venous catheter	(Cavanagh et al., 2014)
ERR085179	*S. haemolyticus*	ERR085179	Clinical	(Cavanagh et al., 2014)
SRR1182430	*S. haemolyticus*	SRR1182430	Clinical	
SRR1182429	*S. haemolyticus*	SRR1182429	Clinical	
SRR1182428	*S. haemolyticus*	SRR1182428	Clinical	
SRR1182432	*S. haemolyticus*	SRR1182432	Clinical	
SRR1182431	*S. haemolyticus*	SRR1182431	Clinical	
ERR085182	*S. haemolyticus*	ERR085182	Clinical	(Cavanagh et al., 2014)
JCSC1435	*S. haemolyticus*	NC_007168	Human	(Takeuchi et al., 2005)
51-30	*S. haemolyticus*	CUDO01000000	Clinical blood	
ERR085171	*S. haemolyticus*	ERR085171	Clinical	(Cavanagh et al., 2014)
25-12	*S. haemolyticus*	CUCI01000000	Clinical blood	
ERR085165	*S. haemolyticus*	ERR085165	Clinical	(Cavanagh et al., 2014)

Isolate name	Species	Accession number	source	Reference
ERR085166	*S. haemolyticus*	ERR085166	Clinical	(Cavanagh et al., 2014)
ERR085174	*S. haemolyticus*	ERR085174	Clinical	(Cavanagh et al., 2014)
ERR085170	*S. haemolyticus*	ERR085170	Clinical	(Cavanagh et al., 2014)
ERR085168	*S. haemolyticus*	E RR085168	Clinical	(Cavanagh et al., 2014)
ERR085173	*S. haemolyticus*	ERR085173	Clinical	(Cavanagh et al., 2014)
ERR085169	*S. haemolyticus*	ERR085169	Clinical	(Cavanagh et al., 2014)
ERR085172	*S. haemolyticus*	ERR085172	Clinical	(Cavanagh et al., 2014)
SH1574	*S. haemolyticus*	LRBM01000000	Clinical eye	(Panda and Singh, 2016)
NW19	*S. haemolyticus*	MRUY01000000	bovine milk	
SH747	*S. haemolyticus*	LRHM01000000	Clinical eye	(Panda and Singh, 2016)
105731	*S. haemolyticus*	CUHI01000000	Catheter	(Cavanagh et al., 2014)
0894-2001-2009	*S. haemolyticus*	QVPX01000000	Umbilical wound	
G811N2B1	*S. haemolyticus*	PGWX01000000	Human nares	
285 SHAE	*S. haemolyticus*	JVMX01000000	Clinical	
FDAARGOS_130	*S. haemolyticus*	LOSE02000000	Clinical	
ERR085180	*S. haemolyticus*	ERR085180	Clinical	(Cavanagh et al., 2014)
708075	*S. haemolyticus*	CUFF01000000	Clinical	
AB	*S. haemolyticus*	CUEN01000000	Human nares	(Cavanagh et al., 2014)
6682	*S. haemolyticus*	CUGF01000000	Clinical blood	(Cavanagh et al., 2014)
C10A	*S. haemolyticus*	JPRW01000000	Clinical sputum	(Chan et al., 2015)
A109N1B1	*S. haemolyticus*	PGWY01000000	Human nares	
83131B	*S. haemolyticus*	CP025396	Clinical	
83131A	*S. haemolyticus*	CP024809	Clinical	

Isolate name	Species	Accession number	source	Reference
ERR085175	*S. haemolyticus*	ERR085175	Clinical	(Cavanagh et al., 2014)
ERR085178	*S. haemolyticus*	ERR085178	Clinical	(Cavanagh et al., 2014)
51-06	*S. haemolyticus*	CUCU01000000	Clinical blood	(Cavanagh et al., 2014)
ERR085183	*S. haemolyticus*	ERR085183	Clinical	(Cavanagh et al., 2014)
51-07	*S. haemolyticus*	CUCV01000000	Clinical blood	(Cavanagh et al., 2014)
ERR085177	*S. haemolyticus*	ERR085177	Clinical	(Cavanagh et al., 2014)
ERR085176	*S. haemolyticus*	ERR085176	Clinical	(Cavanagh et al., 2014)
ERR085181	*S. haemolyticus*	ERR085181	Clinical	(Cavanagh et al., 2014)
6035	*S. haemolyticus*	CUFD01000000	Clinical blood	(Cavanagh et al., 2014)
6249	*S. haemolyticus*	CUFE01000000	Clinical blood	(Cavanagh et al., 2014)
SNUC 3870	*S. hominis*	QXVR01000000	Cow	(Naushad et al., 2016)
SNUC 5336	*S. hominis*	PZHX01000000	Cow	(Naushad et al., 2016)
SNUC 4474	*S. hominis*	QXVP01000000	Cow	(Naushad et al., 2016)
SNUC 2620	*S. hominis*	PZIA01000000	Cow	(Naushad et al., 2016)
SNUC 5852	*S. hominis*	PZHV01000000	Cow	(Naushad et al., 2016)
SNUC 3404	*S. hominis*	PZHY01000000	Cow	(Naushad et al., 2016)
SNUC 2444	*S. hominis*	PZIB01000000	Cow	(Naushad et al., 2016)
SNUC 5746	*S. hominis*	PZHW01000000	Cow	(Naushad et al., 2016)
K1	*S. hominis*	MWPJ01000000	bovine milk	(Wang et al., 2017)
BHG17	*S. hominis*	MPNR01000000	goose droppings	(Naushad et al., 2016)
SNUC 2694	*S. hominis*	PZHZ01000000	Cow	(Lymperopoulou et al., 2017)
H69	*S. hominis*	LVVO01000000	Air from residential area	

Isolate name	Species	Accession number	source	Reference
Hudgins	*S. hominis*	MAYR01000000	Human skin	(Calkins et al., 2016)
J11	*S. hominis*	FBVJ01000000	Human	
C80	*S. hominis*	ACRM01000000	Human	
NCTC_11320	*S. hominis*	PPQE01000000	Human	(Cole et al., 2019)
RE2.10	*S. hominis*	LWJR01000000	rice seed	
UMB0272	*S. hominis*	PKIP01000000	Human	
MMP2	*S. hominis*	LNTW01000000	Ancient permafrost	(Kashuba et al., 2017)
KR	*S. hominis*	NGVM01000000	Kefir seed	
As2	*S. hominis*	LFKR01000000	whole mosquito body	(Hughes et al., 2016)
As3	*S. hominis*	LFKS01000000	whole mosquito body	(Hughes et al., 2016)
As1	*S. hominis*	LFKQ01000000	whole mosquito body	(Hughes et al., 2016)
ZBW5	*S. hominis*	AKGC01000000	Human skin	(Jiang et al., 2012)
CCUG 42399	*S. hominis*	PPQX01000000	Clinical blood	(Cole et al., 2019)
SH04_17	*S. hominis*	PHKJ01000000	Clinical blood	
SH08_17	*S. hominis*	PHKL01000000	Clinical blood	
LRKNS031	*S. hominis*	LXRS01000000	Clinical	
SRR5482196	*S. hominis*	SRR5482196	Clinical blood	
SRR5482200	*S. hominis*	SRR5482200	Clinical blood	
SRR5482295	*S. hominis*	SRR5482295	Clinical blood	
SRR5482291	*S. hominis*	SRR5482291	Clinical blood	
SRR5482198	*S. hominis*	SRR5482198	Clinical blood	
SRR5482201	*S. hominis*	SRR5482201	Clinical blood	
SRR5482203	*S. hominis*	SRR5482203	Clinical blood	

Table 2.3: Accession numbers of isolates used in phylogenetic and pangenome analyses

2.6.8 Accessory genome phylogenetic analysis

The distance of the accessory genome for each sample was determined using the POPpunk pipeline. This software uses variable-length k-mer comparisons to distinguish isolates' divergence in shared sequence (Lees et al., 2019). Firstly, a database is created of all the core and accessory distances between each pair of isolates. Secondly, the database is fitted to a mixture of up to three 2D Gaussians to the distribution of core and accessory distances. The number of mixture components is adjusted for each species to get results that have a low-density score (proportion of edges in the network), high transitivity score and high overall score (Network score based on density and transitivity) (Lees et al., 2019). Accessory genome distance was then plotted on a t-SNE graph with the perplexity (number of close neighbours each point has) adjusted for each species to give the clearest picture of clustering. T-SNE plot was visualised using Microreact, which is free and easily used software on the internet (Argimón et al., 2016).

2.6.9 Pangenome analysis

Pangenome analysis was performed using the high-speed stand-alone pangenome pipeline Roary (version 3.4.2) using the same isolates used in phylogenetic analyses (Page et al., 2015). This tool determines what genes are found in the core genome and what is found in the accessory genome. The Roary pipeline parameter was set to minimum BLASTP percentage identity of 95. The online tool WebMGA was used to assign genes found in the pangenome into their Cluster of Orthologous Groups (COG) family by RPSBLAST which uses a query sequence to search a database of pre-calculated position-specific scoring matrix and reports significant hits in a single pass (Wu et al., 2011). This method is good at identifying protein domains and gene functions within query sequences (Wu et al., 2011).

2.7 Statistical analysis

A Chi-squared test was performed to identify any significant difference in the proportion of multidrug resistant staphylococci and *mecA* gene in isolates recovered from general public settings and public areas in hospitals in East and West London (Campbell, 2007). All percentages were rounded to the nearest tenth. The Chi-squared test was also used to determine the difference in the portion of genes of the COG family that are unique in isolates from general public settings when compared to public areas in hospitals and isolates from East London when compared to isolates from West London. A P value of >0.05 was considered to be significant. The Barnard exact test was performed to identify significance in the proportion of antibiotic resistance genes from WGS sample recovered from general public settings and public areas in hospitals in East and West London (Barnard, 1945). A two-sided P value of >0.05 was considered to be significant.

2.7.1 Hierarchy clustering analysis

Hierarchy clustering of a heatmap for resistance/sensitivity for phenotype and presence absences of genes were created using the R computer language package 'Heatmap.plus' (https://cran.r-project.org/web/packages/heatmap.plus/index.html).

2.7.2 correlation matrix analysis

Pearson correlation was performed on isolates of antibiotic resistant phenotype and genotype data using 'cor' test function in R computer language and plotted using the R computer language package 'corrplot' (https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html).

Chapter 3: Sample collection, species identification of multidrug resistant staphylococci and antibiotic genotype and SCC*mec* element.

3.1 Introduction

Methicillin-resistant *Staphylococcus aureus* (MRSA) and coagulase-negative staphylococci (CoNS) can spread in healthcare and community-associated areas by skin to skin and skin to contaminated surfaces contacts (Conceição et al., 2013; David and Daum, 2010; Xu et al., 2015). Previous studies have shown those non-healthcare associated environments, including recreational beaches, public buses, residential (student) and built-up areas harbour multidrug resistant *S. aureus* (Conceição et al., 2013; Lutz et al., 2014; Roberts et al., 2013). However, studies reporting similar findings for CoNS are fragmentary (Conceição et al., 2013; Mkrtchyan et al., 2013; Roberts et al., 2013; Seng et al., 2017b; Soge et al., 2009; Stepanović et al., 2008; Xu et al., 2015).

The methicillin resistance gene *mecA* is located on a mobile genetic element 'staphylococcal cassette chromosome *mec* (SCC*mec*)'(Oliveira et al., 2002). The *mecA* gene encodes the penicillin-binding protein 2a (PBP2a) that has a low binding affinity to all beta-lactam antibiotics (Stapleton and Taylor, 2002). The SCC*mec* is diverse in its genetic structure and to date, 11 different SCC*mec* types have been characterised. SCC*mec* is determined by the combination of *mec* (A, B, C1, C2, D, E) and the chromosome cassette recombinase (*ccr*) (A1/B1, A2/B2, A3/B3, A4/B4, C1, A5/B3, A1/B6, A1/B3) complexes (Elements (IWG-SCC), 2009; Li et al., 2011; Shore et al., 2011). Different SCC*mec* types have evolved from two different genetic lineages, including hospital-associated and community-associated clones, however, currently, these different lineages can be found both in hospital and community environments (Maree et al., 2007). However, community-associated SCC*mec* types are generally smaller in size compared to their hospital-associated counterparts (2009).

3.2 Method

This chapter describes the overall collection of multidrug-resistance (MDR) staphylococci from high-frequency touched surfaces in public settings in East and West London. These isolates were collected from two sites from East London and West London. One site was from public settings from the community and the other was public areas from Hospitals. The isolates were speciated using Brucker MALDI-TOF MS and validated with the ASTA Tinkerbell MS. Staphylococci antibiotic resistance profile was determined against a panel of 11 different antibiotics as well as the MIC for oxacillin and cefoxitin. From their resistance profile, the proportion of MDR staphylococci from East and West London, public areas in the community and public areas in hospitals can be determined. All isolates that were shown to have *mecA* gene from PCR were whole genome sequenced (WGS). From WGS data the antibiotic resistance genes, SCC*mec* type, the difference between the areas genotype and if the genotype matched with the bacteria phenotype can be determined.

3.3 Results

3.3.1 Sample collection

600 samples collected from November 2016 to September 2017 were screened for multidrug resistance from the general public settings and public areas in hospitals from East and West London (Table 3.1). 224 of these isolates were recovered from East London and 376 from West London. 182 were from general public settings and 418 from public areas in hospitals. 97 samples were recovered from public settings from East London, 85 from public areas in West London. 127 samples were recovered from public areas in hospitals in East London and 291 from Hospitals in West London (Figure 3.1)

	Area			
	ELC	WLC	ELH	WLH
Specific site	washroom door handles Washroom taps Toilet flusher Toilet seat Soap dispensers Door handles Elevator button Bench armrest Escalator rail Stair rail ATM machines Ticket machine Pedestrian crossing buttons	washroom door handles Washroom taps Toilet Flusher Toilet Seat Soap dispensers Door handles Elevator Button Bench armrest Escalator rail Stair rail ATM machines Ticket machine Touch screen TV Public phone Pedestrian crossing buttons	washroom door handles Washroom taps Toilet Flusher Toilet seat Soap dispensers Door handles Elevator button seat armrest Stair hand rail	washroom door handles Washroom taps Toilet flusher Toilet seat Soap dispensers Door handles Elevator button seat armrest Stair hand rail Baby changing area
Number of isolates	97	85	127	291

Table 3.1: Summary of the environmental sites in hospital and community general public areas. ELC= East London Community; WLC=West London Community; ELH=East London Hospital, WLH=West London Hospital. Table in Cave et al., 2019

3.3.2 Multidrug resistant species of staphylococci isolates

281 of the 600 (46.8%) isolates were multidrug resistant staphylococci belonging to 11 species. These included: *S. epidermidis* (n=75); *S. haemolyticus* (n=61); *S. hominis* (n=56)*; S. saprophyticus* (n=24)*; S. warneri* (n=16)*; S. capitas* (n=15)*; S. cohnii* (n=15)*; S. sciuri* (n=9)*, S. aureus* (n=5)*, S. pasteuri* (n=4) *and S. equorum* (n=1). There was a significantly higher proportion of multidrug resistant staphylococci (P=0.0002) recovered from East London (56.7%) compared to those recovered from West London (50.0%) (Table 3.2).

There was a marginally significant difference (P=0.0458) of the proportion of multidrug resistant staphylococcal isolates from public areas in the hospitals to general public settings (49.5% and 40.7% respectively) (Table 3.3).

The most commonly found antibiotic that the staphylococcal isolates were resistant to was penicillin (n=226/80.4%); followed by fusidic acid (n=203/72.2%) erythromycin (n=153/54.5%), amoxicillin (n=78/27.8%); tetracycline (n=74/26.3%); oxacillin (n=70/24.9%); cefoxitin (n=63/22.4%); mupirocin (n=41/14.6%); gentamycin (n=27/9.6%); cefepime (n=20/7.1%), and chloramphenicol (n=11/4.0%).

A hierarchy clustering within a heatmap showed there was no correlation in the species and area they were isolated from to their antibiotic resistance profile (Figure. 3.1). The Chi-square analyses demonstrated that there was a significantly higher proportion of multidrug resistant staphylococci with erythromycin resistance (P= ≤0.0001) and chloramphenicol resistance (P=0.0143) from West London (62.3% and 6.5% respectively) compared to East London (37.0% and 0.8% respectively) (Table 3.1). The opposite was observed for mupirocin where intermediate resistance with a significantly higher

proportion of multidrug resistant staphylococci (P=≤0.0001) was found in East London (19.7%) compared to West London (2.6%) (Table 3.2).

In the general public settings, there was a significantly higher proportion of isolates that had resistance to gentamycin (P=0.00162) and tetracycline (P=0.0211) (16.2% and 36.5% respectively) compared to public areas in hospitals (36.5% and 22.7% respectively) (Table 3.2). In contrast, a significantly higher proportion of multidrug resistant staphylococci (P=0.0143) found in public areas in hospitals (26.1%) were resistant to cefoxitin compared to general public settings (12.2%).

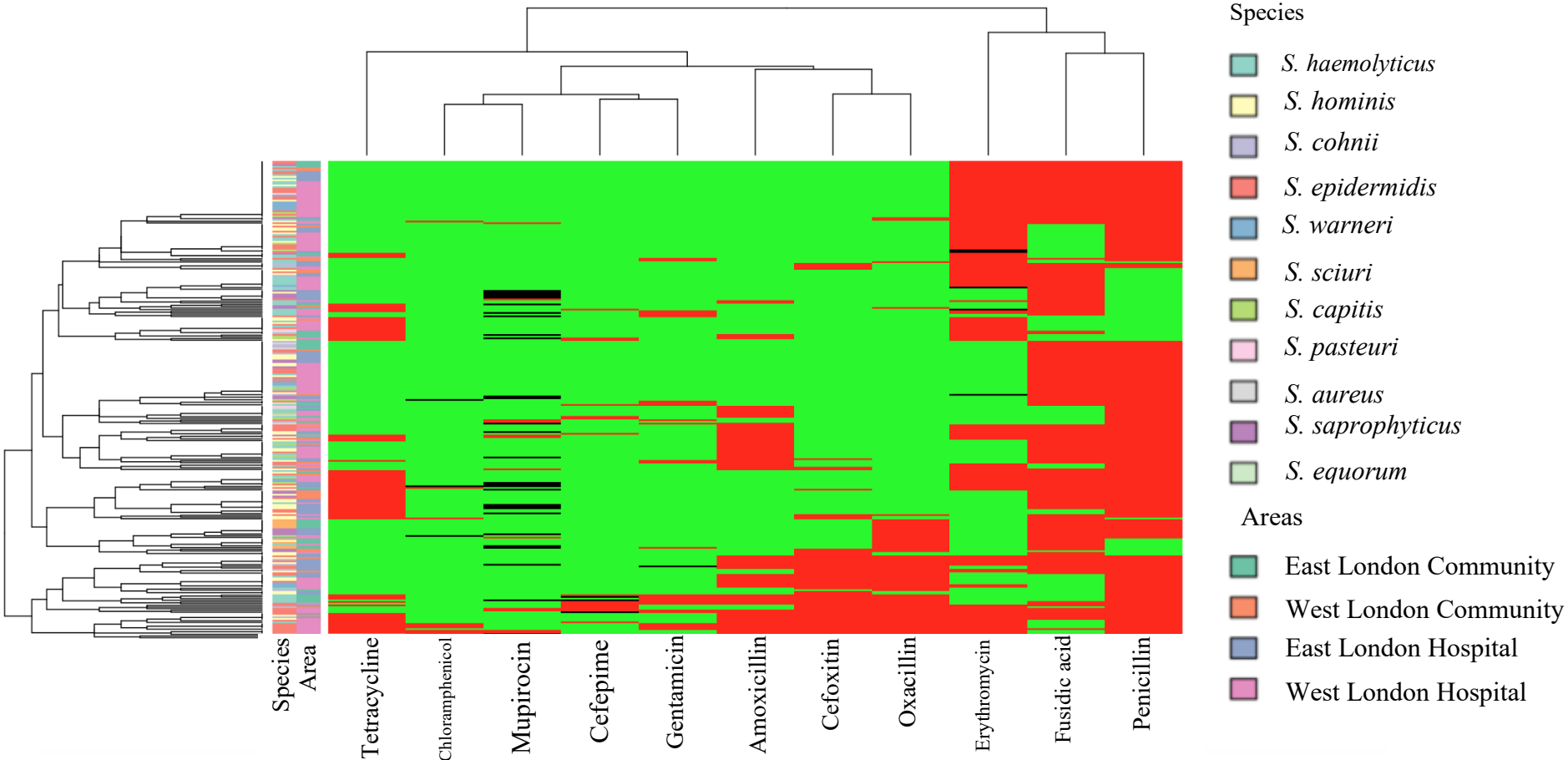


Figure 3.1: **Hierarchical clustering heatmap showing there was no clustering of the antibiotic resistance profile of isolates in comparison with the species and area they were isolated from.** Red tile indicates resistance, black tiles represent intermediate resistance and green represent sensitive. Figure in Cave et al. 2019

	East London Total number of isolates screened (N = 224)		West London Total number of isolates screened (N=376)		Chi-Square test		
	N	% of total number of isolates screened	N	% of total number of isolates screened	% Difference	X^2	P value
Multidrug resistant staphylococci	127	56.7	154	41.0	15.7	13.944	0.0002
mecA positive	24	10.7	27	7.18	3.5	2.246	0.134
	N	% MR staphylococci	N	% of MR staphylococci	% Difference	X^2	P value
Oxacillin	38	29.9	32	20.8	9.1	3.097	0.0784
Gentamicin R	13	10.2	13	8.4	1.8	0.268	0.6049
Gentamicin I	1	0.8	0	0	0.8	1.217	0.27
Mupirocin R	4	3.2	8	5.2	2.1	0.706	0.4006
Mupirocin I	25	19.7	4	2.6	17.1	21.87	<0.001
Amoxicillin	33	26.0	45	29.2	3.3	0.363	0.5468
Erythromycin R	47	37.0	96	62.3	25.3	17.80	<0.001
Erythromycin I	1	0.8	5	3.3	2.5	2.006	0.1567
Tetracycline	36	28.4	38	24.7	3.7	0.481	0.4878
Cefoxitin	29	22.8	34	22.1	0.8	0.022	0.8809
Cefepime R	7	5.5	10	6.5	1.0	0.117	0.7321
Cefepime I	2	1.6	1	0.7	0.9	0.557	0.4556
Fusidic acid	97	76.4	106	68.8	7.6	1.971	0.1603
Penicillin	102	80.3	124	80.5	0.2	0.002	0.9648
Chloramphenicol R	1	0.8	10	6.5	5.7	5.992	0.0144
Chloramphenicol I	1	0.8	2	1.3	0.5	0.17	0.6997

Table 3.2: The proportion of multidrug resistant staphylococci and *mecA*$^{+}$ isolates compared with the number of isolates screened in East and West London and the proportion of antibiotics they were resistant compared with the number of multidrug resistant staphylococci from East and West London. All chi-squared test was performed with 1 degree of freedom. R= resistance; I= intermediate resistance; MR= multidrug resistant. Table in Cave et al., 2019

	General public settings Total number of isolates screened (n=182)		Public areas in hospitals Total number of isolates screened (n=418)		Chi-Square test		
	n	% of the total number of isolates screened	n	% of the total number of isolates screened	% Difference	X^2	P value
Multidrug resistant staphylococci	74	40.7	207	49.5	8.9	3.991	0.0458
mecA positive	14	7.7	33	7.9	0.2	0.007	0.9332
Antibiotic resistance	N	% MR staphylococci	N	% MR staphylococci	% Difference	X^2	P value
Oxacillin	24	32.4	46	22.2	10.2	3.097	0.0784
Gentamicin R	12	16.2	14	6.8	9.5	5.79	0.0161
Gentamicin I	0	0.0	1	0.5	0.5	0.355	0.5512
Mupirocin R	2	2.7	10	4.8	2.1	0.603	0.603
Mupirocin I	6	8.1	23	11.1	3.0	0.528	0.4674
Amoxicillin	18	24.3	60	29.0	4.7	0.591	0.4421
Erythromycin R	33	44.6	110	53.1	8.6	1.589	0.2075
Erythromycin I	1	1.4	5	2.4	1.1	0.297	0.5856
Tetracycline	27	36.5	47	22.7	13.8	5.316	0.0211
Cefoxitin	9	12.2	54	26.1	13.9	6.06	0.0138
Cefepime R	7	9.5	10	4.8	4.6	2.049	0.1523
Cefepime I	2	2.7	1	0.5	2.2	2.542	0.1109
Fusidic acid	54	73.0	149	72.0	1.	0.027	0.8706
Penicillin	56	75.7	170	82.1	6.5	1.436	0.2308
Chloramphenicol R	1	1.4	10	4.8	3.5	1.749	0.186
Chloramphenicol I	0	0.00	3	1.5	1.5	1.081	0.2985

Table 3.3: The proportion of multidrug resistant staphylococci and *mecA*$^+$ isolates compared with the number of isolates screened in general public settings and hospitals; the proportion of antibiotics that were resistant compared to the number of multidrug resistant staphylococci from general

public settings and hospitals. All chi-squared test was performed with 1 degree of freedom. R= resistance; I= intermediate resistance; MR= multidrug resistant . Table in Cave et al. 2019.

3.3.3 MALDI-TOF validation

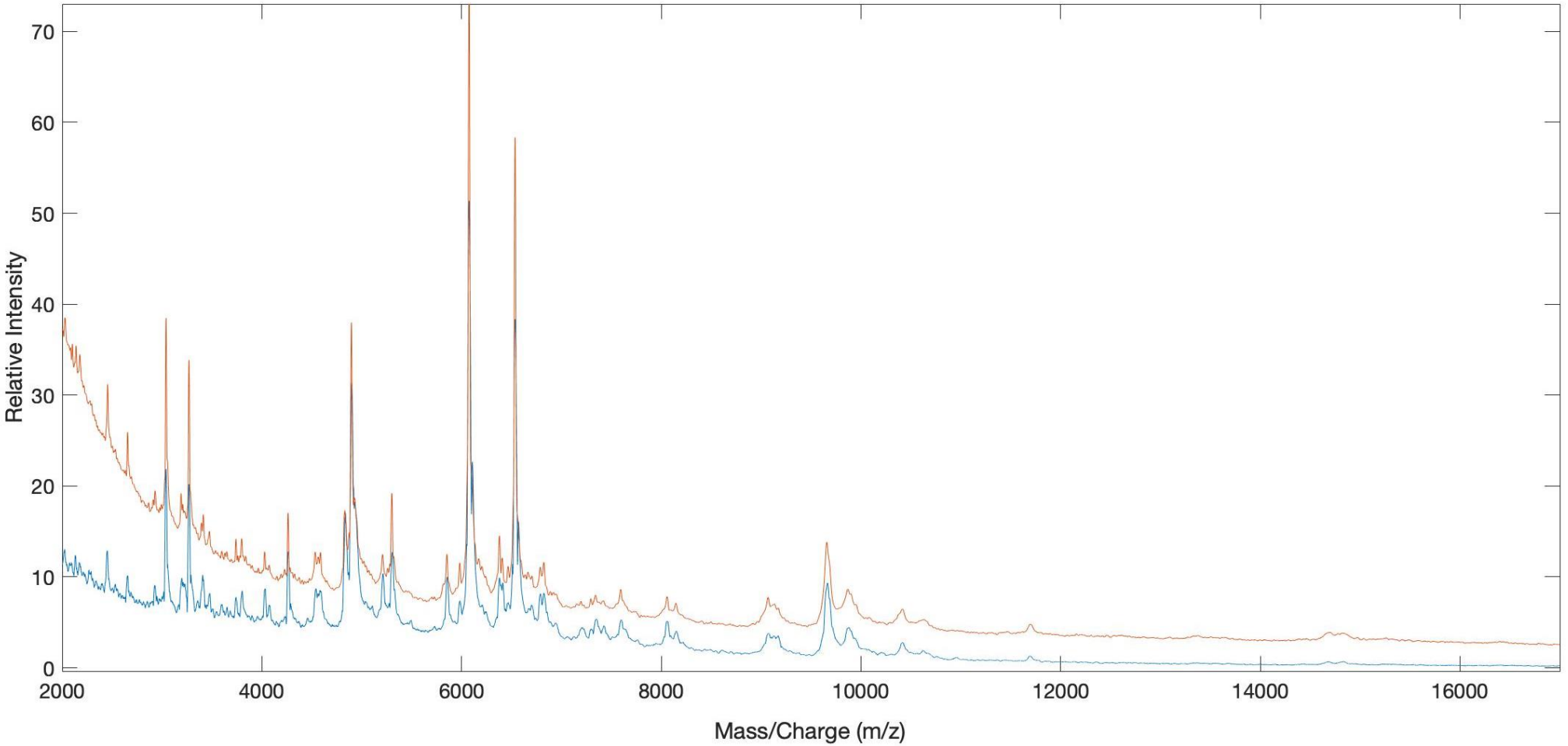
The accuracy of these results was checked with Brucker instruments by testing 89 MDR isolates from 10 different species. This includes *S. aureus* (n=5); *S. capitis* (n=10)*; S. cohnii* (n=10); *S. epidermidis* (n=11); *S. haemolyticus* (n=11); *S. hominis* (n=11) ; *S. pasteuri* (n=2); *S. sciuri* (n=6); *S. saprophyticus* (n=12) and *S. warneri* (n=11) (Table 3.4). Brucker's Autoflex was able to predict 7 species correctly 100% of the time whereas ASTA's Tinkerbell predicted 6 species correctly 100% of the time. Of the species, they predicted correctly 100% of the time only 3 species for each instrument was predicted at high confidence 100% of the time. These were *S. epidermidis* and *S. saprophyticus* for both instruments; *S. saprophyticus* for Brucker's Autoflex and *S. capitis* for ASTA's Tinkerbell. Only *S. cohnii* (3.3%) for Brucker Autoflex and *S. aureus* (37.5%) and *S. sciuri* (0%) were predicted at low confidence.

Species	No isolate	Bruker		ASTA	
		% High confidence	% Correct ID	% High confidence	% Correct ID
S. aureus	5	95.8	95.8	37.5	91.0
S. capitis	10	96.7	100	100	100
S. cohnii	10	3.3	86.7	80.0	80.0
S. epidermidis	11	100	100	100	100
S. haemolyticus	11	90.9	100	91	100
S. hominis	11	100	100	97	100
S. pasteuri	2	83.3	100	66.7	100
S. sciuri	6	88.9	100	0	55.6
S. saprophyticus	12	100	100	100	100
S. warneri	11	91.0	100	94	97.0

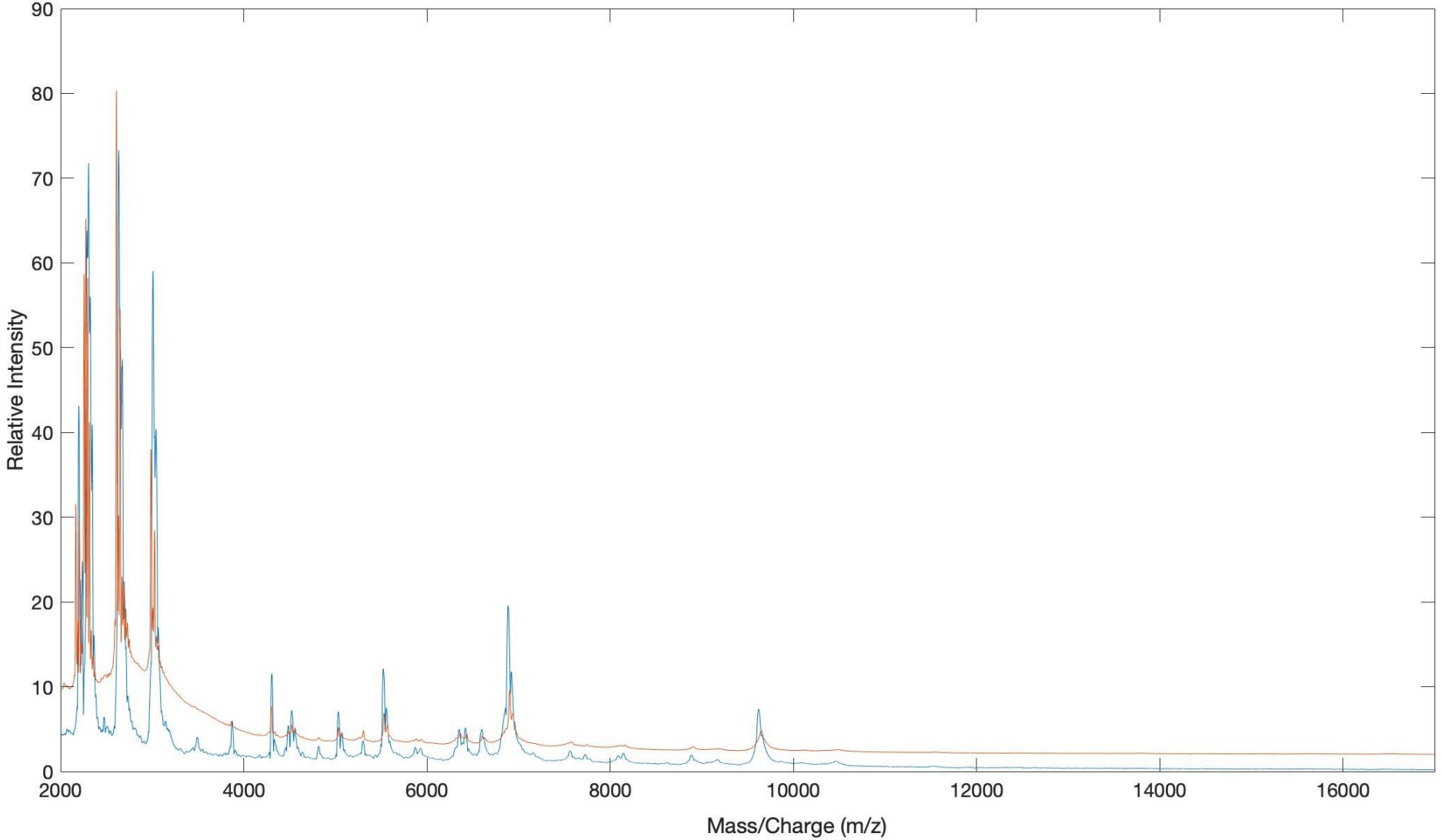
Table 3.4: Percentage of 92 environmental staphylococci isolates which were correctly identified by two MALDI-TOF instruments.

The two instruments' mass spectrum was compared with all isolates from the environment. There were comparable mass ions peaks for the same isolates on both instruments. These also included the *S. cohnii* isolates which were predicted correctly at high confidence by the ASTA Tinkerbell instrument but were misidentified or identified at low confidence on Bruker's Autoflex and vice versa with the *S. aureus* and *S. sciuri* isolates (Figure 3.2). The only species which had 100% correct ID to high confidence for both instruments was *S. epidermidis* and *S. saprophyticus.*

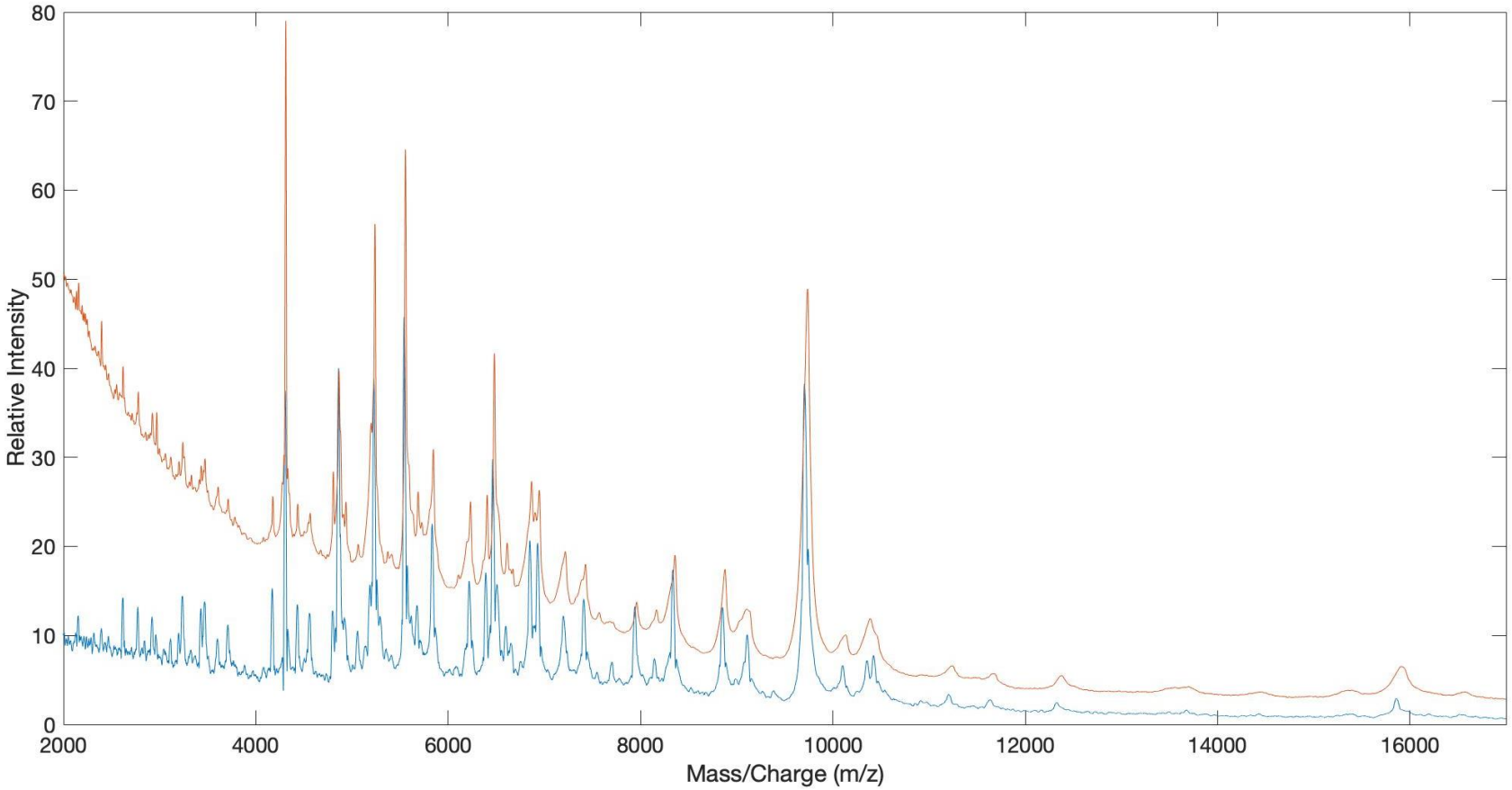
A


Relative Intensity
0
10
20
30
40
50
60
70
2000
4000
6000
8000
10000
12000
14000
16000
Mass/Charge (m/z)

B


Relative Intensity
Mass/Charge (m/z)
0
10
20
30
40
50
60
70
80
90
2000
4000
6000
8000
10000
12000
14000
16000

C



on the other MS platform which had similar mass ion peaks.

A= *S. cohnii*; B= *S. aureus*, C= *S. sciuri.* Blueline Brucker Autoflex, Redline Asta Tinkerbell

3.3.4 Detection of *mecA* gene

The *mecA* gene was identified in 49 (8.2%) isolates. There was no significant difference in the proportion of the *mecA* gene determined in isolates recovered from East London (10.7%) compared to those recovered from West London (7.2%) (P=0.1340), the general public settings (7.7%) and public areas in hospitals (7.2%) (P = 0.9332). Of the isolates that were *mecA*$^{+}$, 44 (62.9%) were oxacillin resistant, whereas 43 (68.3%) isolates were cefoxitin resistant. Five isolates that were *mecA*$^{+}$ were sensitive to oxacillin and 6 *mecA*$^{+}$ isolates (all belonging to the *S. sciuri* species) were sensitive to cefoxitin.

3.3.5 Determination of MICs for oxacillin and cefoxitin

The MICs for oxacillin and cefoxitin were determined for 49 isolates that carried the *mecA* gene (Table 3.5). Although all isolates were *mecA*$^{+}$, only 44 CoNS isolates had MIC above the resistance breakpoints, according to CSLI, 2017. Five isolates, including *S. hominis* 372, 385, 387; *S. epidermidis* 465 and *S. haemolyticus* 361 that were *mecA*$^{+}$, were phenotypically oxacillin sensitive. However, all five isolates were resistant to cefoxitin by zone diffusion assay. These isolates were recovered from public areas in hospitals. Neither CLSI nor BSAC recommend MIC standards for recoding cefoxitin resistance. Nevertheless, 42 out of 43 isolates in this study had MIC values of >1.5 μg/ml and were resistant to cefoxitin as shown by a disc diffusion assay.

Isolate	species	Areas in London	Oxa	Gen	Mup	Amx	Erm	Tet	Fox	Fep	Fua	Pen	Chl	Oxa MIC (µg/ml)	Fox MIC (µg/ml)
1	*S. haemolyticus*	ELC	R	R	S	R	S	R	R	R	S	R	S	3	4
27	*S. sciuri*	ELC	R	S	S	S	S	S	S	S	R	R	S	0.5	0.75
33	*S. sciuri*	ELC	R	S	S	S	S	S	S	S	R	R	S	0.5	1
59	*S. sciuri*	ELC	R	S	S	S	S	S	S	S	R	S	S	0.75	1
74	*S. sciuri*	ELC	R	S	S	S	S	S	S	S	R	S	S	0.5	1
75	*S. sciuri*	ELC	R	R	I	S	S	S	S	S	R	S	S	0.75	1
93	*S. haemolyticus*	ELC	R	R	I	R	S	S	R	I	S	R	S	2	4
99	*S. haemolyticus*	ELC	R	R	S	R	S	R	R	S	S	R	S	3	4
105	*S. haemolyticus*	ELC	R	R	S	R	S	R	R	R	R	R	S	2	4
109	*S. sciuri*	ELC	R	S	S	S	S	S	S	S	R	R	S	1	1
207	*S. hominis*	WLC	R	S	S	S	R	S	R	S	R	R	S	0.5	6
208	*S. hominis*	WLC	R	S	S	R	R	R	R	S	R	R	S	2	6
209	*S. hominis*	WLC	R	S	S	R	R	S	R	S	R	R	S	1.5	6
211	*S. cohnii*	WLC	R	S	S	S	R	S	R	R	S	R	S	4	4
321	*S. epidermidis*	ELH	R	S	S	R	R	S	R	S	R	R	S	0.75	3
327	*S. epidermidis*	ELH	R	I	S	R	S	S	R	S	R	R	S	0.75	2
329	*S. epidermidis*	ELH	R	S	S	R	R	R	R	S	R	R	S	0.75	8

Isolate	species	Areas in London	Oxa	Gen	Mup	Amx	Erm	Tet	Fox	Fep	Fua	Pen	Chl	Oxa MIC (µg/ml)	Fox MIC (µg/ml)
343	*S. cohnii*	ELH	R	S	S	R	R	R	R	S	R	R	S	1.5	12
349	*S. cohnii*	ELH	R	S	S	R	R	S	R	S	R	R	S	1.5	12
355	*S. epidermidis*	ELH	R	S	S	R	R	S	R	S	R	R	S	0.5	3
361	*S. haemolyticus*	ELH	S	S	S	R	S	S	R	S	R	R	S	0.38	4
372	*S. hominis*	ELH	S	S	S	S	S	S	R	S	S	R	S	0.25	6
373	*S. haemolyticus*	ELH	R	S	S	R	S	S	R	S	S	R	S	1	8
385	*S. hominis*	ELH	S	S	S	S	S	S	R	S	S	R	S	0.125	1.5
386	*S. hominis*	ELH	R	S	S	R	R	S	R	S	R	R	S	4	0.38
387	*S. hominis*	ELH	S	S	R	R	R	S	R	S	R	R	S	0.064	16
407	*S. epidermidis*	ELH	R	S	S	R	S	S	R	S	R	R	S	0.5	4
435	*S. epidermidis*	WLH	R	R	S	R	S	S	R	R	R	R	S	1	6
436	*S. epidermidis*	WLH	R	S	S	R	R	R	R	S	S	R	S	1.5	8
445	*S. haemolyticus*	WLH	R	S	I	R	R	S	R	S	S	R	R	4	4
465	*S. epidermidis*	WLH	S	S	S	R	R	R	R	R	R	R	R	0.38	2
475	*S. epidermidis*	WLH	R	S	R	S	R	S	R	R	R	R	S	2	12
479	*S. hominis*	WLH	R	S	S	R	R	R	R	S	R	R	S	1.5	16
492	*S. haemolyticus*	WLH	R	S	S	S	S	S	R	S	S	R	S	0.75	8
506	*S. haemolyticus*	WLH	R	S	S	R	R	R	R	R	S	R	S	4	12

Isolate	species	Areas in London	Oxa	Gen	Mup	Amx	Erm	Tet	Fox	Fep	Fua	Pen	Chl	Oxa MIC (µg/ml)	Fox MIC (µg/ml)
538	*S. haemolyticus*	WLH	R	S	S	R	I	S	R	R	R	R	R	0.5	6
620	*S. hominis*	WLH	R	S	S	R	S	S	R	S	S	R	S	3	16
623	*S. hominis*	WLH	R	S	S	R	S	S	R	S	S	R	S	2	24
631	*S. epidermidis*	WLH	R	R	S	R	R	S	R	I	R	R	S	3	16
664	*S. epidermidis*	WLH	R	S	S	R	S	S	R	S	S	R	S	2	6
673	*S. epidermidis*	WLH	R	R	S	R	R	R	R	S	R	R	S	4	3
699	*S. warneri*	WLH	R	S	S	R	S	S	R	S	S	R	S	3	8
700	*S. warneri*	WLH	R	S	S	R	R	S	R	S	S	R	S	4	6
702	*S. warneri*	WLH	R	S	S	R	R	S	R	S	S	R	S	2	12
711	*S. epidermidis*	WLH	R	R	S	R	R	R	R	S	S	R	R	12	24
712	*S. epidermidis*	WLH	R	R	S	R	R	R	R	S	S	R	R	12	24
713	*S. epidermidis*	WLH	R	R	S	R	R	R	R	S	S	R	R	256	12
715	*S. epidermidis*	WLH	R	S	R	R	R	R	R	S	S	R	R	256	12
716	*S. epidermidis*	WLH	R	S	R	R	R	R	R	S	S	R	R	256	12

Table 3.5: The antibiotic resistance profile of 49 *mecA*$^{+}$ isolates recovered from public areas in hospitals and general public settings.

R = resistant; I = intermediate resistance, S = sensitive; Oxa = oxacillin; Gen = gentamycin; Mup = mupirocin; Amx = amoxicillin; Erm = erythromycin; Tet = tetracycline; Fox= cefoxitin; Fep = cefepime; Fua= fusidic acid; Pen= penicillin; Chl= chloramphenicol ELC= East London Community; WLC= West London Community; ELH= East London Hospital; WLH= West London hospital. Table in Cave et al. 2019

3.3.6 *De novo* assembly statistics of whole genome sequencing data

The sequenced reads were assembled for the 49 *mecA*$^{+}$ staphylococcal isolates that were WGS to produce a draft genome (Table 3.6). Contigs less then 500bp were filtered out. The assembly ranged from 14-99 contigs; N50 of 67,338-1,568,201and L50 1-14. Genome size and GC content for *S. epidermidis* range from 2,423,410 to 2,632,209 with a GC content 31.3 to 32.1%. For *S. haemolyticus* 2,377,188 to 2,597,964bp with a GC content of 32.6 to 32.7%. For *S. hominis* 2,114,977 to 2,243,945bp with a GC content of 31.27 to 31.4%. For *S. cohnii* 2,672,915 to 2,710,971bp with a GC content of 32.37 to 32.5%. *S. warneri* 2,408,240 to 2,408,986bp with GC content 32.6%; and *S. sciuri* 2,780,223 to 2,784,151bp with a GC content of 32.5 to 32.6%. The number of contigs in *S. epidermidis* ranged 34 to 99 with a mean of 47.6% of the contigs larger than 10,000bp and 24.0% of the contigs bigger than 50,000bp. For *S. haemolyticus* the number of contigs ranged from 65 to 96 with a mean of 61.5% of the contigs larger than 10,000bp and 24.1% of the contigs larger than 50,000bp. *S. hominis* isolates have a number of contig range from 13 to 81 with a mean of 64.4% of isolates contig larger than 10,000bp and 34.8% of the contigs larger than 50,000bp. *S. cohnii* number of contigs range from 32 and 48 with a mean of 51.6% of the isolates contig larger than 10,000bp and 31.8% of isolates contig is larger than 500,000bp. *S. warneri* number of contigs range from 17 to 18 with a mean of 50% of the contigs larger than 10,000bp and 30.8% of the isolates contig is larger than 50,000bp. *S. sciuri* number of contigs range from 14 to 22 with a mean of 47.27% of the contigs larger than 10,000bp and 30.9% of the contigs larger than 50,000bp.

Sample	Species	# contigs	# contigs (>= 10000 bp)	# contigs (>= 25000 bp)	# contigs (>= 50000 bp)	Largest contig	Total length	N50	L50	GC (%)
1	*S. haemolyticus*	71	44	28	21	198,215	2,426,191	76,937	11	32.7
27	*S. sciuri*	22	10	8	7	1,460,882	2,781,043	1,460,882	1	32.5
33	*S. sciuri*	20	9	8	6	1,462,178	2,784,151	1,462,178	1	32.5
59	*S. sciuri*	17	8	7	5	1,567,438	2,780,228	1,567,438	1	32.5
74	*S. sciuri*	16	9	8	6	1,462,149	2,780,223	1,462,149	1	32.5
75	*S. sciuri*	14	8	7	5	1,568,201	2,779,680	1,568,201	1	32.5
93	*S. haemolyticus*	65	37	25	13	391,966	2,435,134	87,864	7	32.7
99	*S. haemolyticus*	80	47	31	22	139,290	2,597,964	88,243	12	32.7
105	*S. haemolyticus*	76	45	30	20	198,215	2,429,313	76,937	11	32.7
109	*S. sciuri*	21	8	7	5	1,567,274	2,783,685	1,567,274	1	32.6
207	*S. hominis*	45	31	23	15	309,626	2,243,945	107,022	7	31.4
208	*S. hominis*	39	23	18	17	326,099	2,250,725	122,411	6	31.3
209	*S. hominis*	48	32	23	15	307,785	2,225,890	106,764	7	31.4
211	*S. cohnii*	32	24	23	16	375,488	2,710,971	194,587	5	32.4
321	*S. epidermidis*	41	26	19	15	296,717	2,573,368	145,078	6	32.0
327	*S. epidermidis*	39	24	18	14	365,826	2,573,135	168,921	6	32.0
329	*S. epidermidis*	44	28	22	16	253,615	2,570,296	141,020	7	32.0
343	*S. cohnii*	48	21	18	12	519,772	2,689,027	274,787	4	32.4
349	*S. cohnii*	46	20	16	12	526,250	2,672,915	296,280	4	32.5
355	*S. epidermidis*	40	24	20	13	434,249	2,499,607	159,061	5	31.9
361	*S. haemolyticus*	96	55	29	14	243,763	2,444,366	52,518	12	32.7
372	*S. hominis*	34	21	15	12	466,679	2,171,866	159,266	4	31.4
373	*S. haemolyticus*	69	42	28	18	251332	2,377,188	78,596	11	32.7
385	*S. hominis*	13	10	8	7	1016183	2,170,655	522,044	2	31.4
386	*S. hominis*	81	31	18	9	466504	2,214,036	133,680	5	31.3
387	*S. hominis*	41	25	18	15	317973	2,231,903	135,176	5	31.3

407	*S. epidermidis*	34	23	20	15	356468	2,423,410	143,743	6	32.1
435	*S. epidermidis*	52	29	24	17	254787	2,632,209	120,838	7	31.8
436	*S. epidermidis*	51	29	24	18	285291	2,463,964	105,346	8	32.0
445	*S. haemolyticus*	70	50	31	14	211088	2,496,057	65,245	11	33.0
465	*S. epidermidis*	39	22	18	13	517090	2,485,514	226,829	4	32.0
475	*S. epidermidis*	45	29	20	15	518475	2,506,454	120,888	6	32.0
479	*S. hominis*	41	23	16	14	312950	2,248,765	154,197	5	31.3
492	*S. haemolyticus*	38	23	19	14	257061	2,339,728	158,248	6	32.7
506	*S. haemolyticus*	69	43	26	16	286475	2,457,490	87,637	10	32.6
538	*S. haemolyticus*	67	45	27	17	228797	2,484,453	77,667	10	32.6
620	*S. hominis*	43	24	18	14	411659	2,114,977	140,509	5	31.4
623	*S. hominis*	32	21	15	12	466681	2,172,479	159,266	4	31.4
631	*S. epidermidis*	60	36	24	16	326412	2,461,843	103,579	8	32.0
664	*S. epidermidis*	42	24	20	15	518823	2,522,035	180,767	5	32.0
673	*S. epidermidis*	63	39	30	16	209796	2,483,246	95,655	9	32.0
699	*S. warneri*	17	8	8	6	1271043	2,408,986	1,271,043	1	32.6
700	*S. warneri*	18	10	9	5	1194182	2,408,574	611,334	2	32.6
702	*S. warneri*	17	8	7	5	1270917	2,408,240	1,270,917	1	32.6
711	*S. epidermidis*	61	34	28	20	204104	2,594,970	93,620	9	31.8
712	*S. epidermidis*	61	34	28	20	204104	2,594,586	93,620	9	31.8
713	*S. epidermidis*	62	34	28	20	204104	2,594,914	93,620	9	31.8
715	*S. epidermidis*	67	34	28	20	204104	2,569,288	93,539	9	31.8
716	*S. epidermidis*	99	48	34	18	177708	25,99,974	67,338	14	31.7

Table 3.6: Genome assembly statistics of isolates recovered from general public settings in East and West London.

3.3.7 Prevalence of antibiotic genes from WGS data

The *mecA* gene was found in 43 out of the 49 isolates that were whole genome sequenced. Of these all *S. sciuri* isolates did not carry the *mecA* gene. Instead, they carried the *mecA1* gene, which had only 84.4% homology to *mecA* gene.

Apart from *mecA,* 24 other antibiotic resistance genes were detected in 43 isolates. *BlaZ* was most commonly found resistance gene with 39 isolates (90.7%) followed by *qacA/B* with 22 (51.2%); *dfrC* with 18 (41.9%), *norA* and *ant(4')-lb* with 17 (39.5%); *AAC(6')-Ie-APH(2'')-Ia* with 15 (34.9%), *fusB* with 14 (32.6%), *msrA* with 13 (30.2%), *ermC* with 12 (27.9%), *mphC* with 9 (27.6%), *tetK* 8 (18.6%), *mupA* with 7 (16.3%), *cat* with 6 (14.0%), *dfrG* with 5 (11.63%), *mgrA* with 5 (9%), *lnuA* with 4 (9.3%), *fusC* and *aph3-IIIa* with 3 (7.0%) and *sat4A*, *vgaA*, *vatB* which were all found in 1 isolate (2.3%).

From these 43 isolates, 3 (7.0%) isolates had two antibiotic resistance genes; 3 (7.0%) had three antibiotic resistance genes; 7 (16.9%) had four antibiotic resistance genes , 2 (4.7%) had five antibiotic resistance genes, 7 (16.3%) had six antibiotic resistance genes, 2 (4.7%) had seven antibiotic resistance genes, 3 (7.0%) had eight antibiotic resistance genes, 6 (14.0%) had nine antibiotic resistance genes and 5 (11.6%) had ten antibiotic resistance genes.

A hierarchy clustering within a heatmap of the *mecA*$^{+}$ isolates resistance gene profile has shown a clustering of *S. epidermidis* isolates except for sample 407 and 465 as well as all *S. warneri* isolates and *S. haemolyticus* from East London community (Figure 3.3). Interestingly, all *S. epidermidis* isolates had the *norA* and *dfrC* genes.

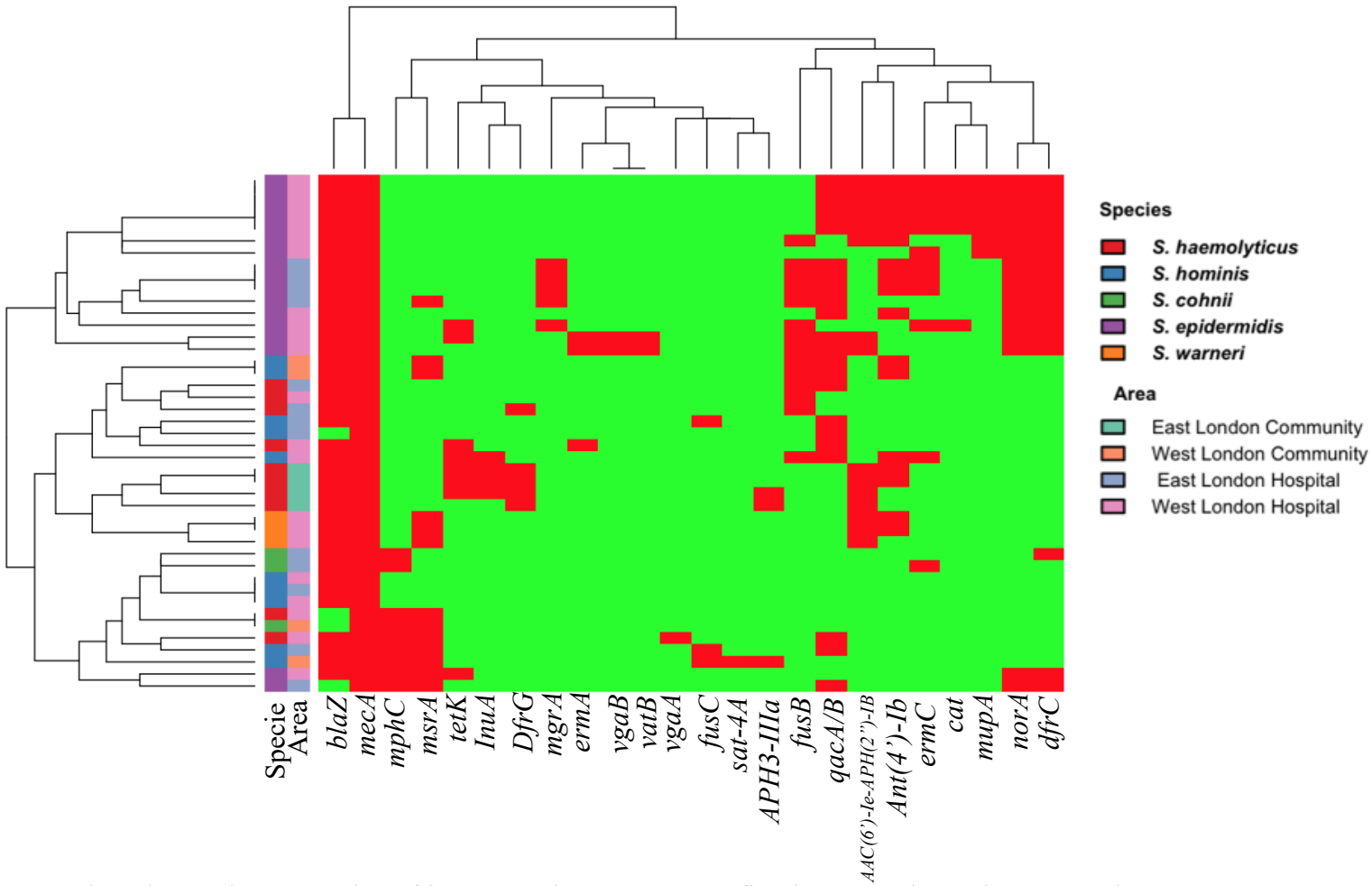


Figure 3.3: Heatmap showing hierarchical clustering of isolates resistance gene profiles in comparison with the species and area they were isolated from.

Red tile indicates the presence of antibiotic resistance genes; green tile absence of resistance gene. Figure in Cave et al., 2019

Barnard's Exact test analysis showed there was a significantly higher proportion of isolates with the *dfrG* gene (P=0.0054) in East London (29.4%) compared to West London (0%) (Table 3.7). There was a significantly higher proportion of isolates with the *cat* (P=0.0419) and *mup* gene (P=0.0238) in West London (23.1% and 26.9% respectively) and compared to East London (both 0%).

For general public settings there was significantly higher proportion of antibiotics *aph2-IIIa* (P=0.0024), *lnuA* (P= 0.0116) and *dfrG* (P=0.0031) (25%, 37.5% and 50% respectively) compared to public areas in hospitals (0%, 0% and 2.86% respectively) (Table 3.8). The opposite was observed with isolates carrying the *dfrC* (P=0.0238), and *norA* gene (P=0.0238) with a significantly higher proportion found in public areas in hospitals (51.4% and 48.6% respectively) compared with general public settings (both 0%).

	East London Total number of isolates WGS (N=17)		West London Total number of isolate WGS (N=26)		Barnard Exact Test	
Antibiotic resistance genes	N	% of total number of isolates WGS	N	% of total number of isolates WGS	Difference	P value
blaZ	15	88.2	24	92.3	4.1	0.7224
tetK	3	17.7	11	19.2	1.6	0.9565
ant(4')-lb	5	29.4	12	46.2	16.7	0.3766
AAC(6')-Ie-APH(2'')-	4	23.5	11	42.3	18.8	0.2291
Ia	2	11.8	1	3.9	8.0	0.4623
aph3-IIIa	3	17.7	1	3.9	13.8	0.1749
lnuA	5	29.4	0	0	29.4	0.0054
DfrG	6	35.3	12	46.2	10.9	0.7546
DfrC	6	35.3	8	30.8	4.5	0.6665
fusB	2	11.8	1	3.9	7.9	0.4623
fusC	9	52.9	13	50	2.9	0.9565
qac	3	17.7	10	38.5	20.8	0.2175
msrA	0	0	1	3.9	3.9	0.4872
Sat4A	3	17.7	5	19.2	1.6	0.9565
mphC	5	29.4	12	46.2	16.7	0.3766
norA	4	23.5	1	3.9	19.7	0.0657
mgrA	0	0	3	11.5	11.5	0.2065
ermA	4	23.5	8	30.8	7.2	0.7224
ermC	0	0	7	26.9	26.9	0.0238
mupA	0	0	6	23.1	23.1	0.0419
cat	0	0	1	3.9	3.9	0.4872
vgaA	0	0	2	7.7	7.7	0.3766
vgaB	0	0	2	7.7	7.7	0.3766
vatB						

Table 3.7: The proportion of antibiotic resistance genes in isolates recovered from East and West London that possessed the *mecA* gene. Table in Cave et al. 2019

	General public settings Total number of isolatws WGS (N=8)		Public areas in hospitals Total number of isolate WGS (N=35)		Barnard Exact Test	
Antibiotic resistance genes	N	% of total number of isolates WGS	N	% of total number of isolates WGS	Difference	P value
blaZ	7	87.5	32	91.4	3.9	1
tetK	3	37.5	5	14.3	23.2	0.1810
ant(4')-lb	4	50	12	37.1	12.9	0.8026
AAC(6')-Ie-APH(2'')-Ia	4	50	11	31.4	18.6	0.4519
aph3-IIIa	3	37.5	0	0	37.5	0.0024
lnuA	3	37.5	1	2.9	34.6	0.0116
DfrG	4	50	18	2.9	47.1	0.0031
DfrC	0	0	12	51.4	51.4	0.0238
fusB	6	25	2	25.7	0.7	1.0000
fusC	1	12.5	20	5.7	6.8	0.8026
qacB	2	25	9	57.1	32.1	0.1808
msrA	4	50	0	25.7	24.3	0.2078
Sat4A	1	12.5	7	0	12.5	0.1664
mphC	2	25	17	20	5.0	1.0000
norA	0	0	5	48.6	48.6	0.0238
mgrA	0	0	3	14.3	14.3	0.3686
ermA	0	0	12	8.6	8.6	0.5992
ermC	0	0	7	34.3	34.3	0.1664
mupA	0	0	6	20	20.0	0.1945
cat	0	0	1	17.1	17.1	0.2668
vgaA	0	0	2	2.9	2.9	0.8160
vgaB	0	0	2	5.7	5.7	0.8026
vatB	0	0	2	5.7	5.7	0.8026

Table 3.8:The proportion of antibiotic resistance genes in isolates recovered general public settings and public areas in hospitals that possessed the *mecA* gene. Table in Cave et al. 2019

3.3.8 Correlation of antibiotic resistance phenotype and genotype using WGS analysis

Antibiotic resistance phenotype and genotype were compared to determine whether the phenotype of the isolates correlates with the genes responsible for resistance to particular antibiotics to which they were resistant (Table 3.9). The data showed that not all phenotypes correlated with the predicted genotype. For better visualisation a Pearson correlation was performed on all isolates that were WGS sequenced that were shown to have known antibiotic resistance genes (Figure 3.4). *MecA* gene, cefoxitin and penicillin resistant phenotypes were removed from the Pearson correlation analysis as all isolates were shown to have these traits. There was a strong correlation (p=>0.05) for gentamicin and aminoglycoside resistance gene *aac(6')-le-aph(2")-la* (r=0.73); amoxicillin and the beta-lactam resistant *blaZ* gene (r=0.56), tetracycline and *tetK* (r=0.62), fusidic acid *fusB* gene (r=0.51) and chloramphenicol and *cat* gene (r=0.67). mupirocin had a weak correlation to *mupA* (r=0.456) and erythromycin a weak correlation to *ermC* (r=0.38) and *ermA* (r=0.29).

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
1	*S. haemolyticus*	Oxacillin Gentamicin Amoxicillin Tetracycline Cefoxitin Cefepime Penicillin	Penam Aminoglycoside Penam Tetracycline Cephalosporin cephalosporin Penam	*mecA* *BlaZ* *DfrG* *TetK* *Ant4-IB* *Lnu* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Diaminopyrimidine Tetracycline Aminoglycoside Lincosamide Aminoglycoside
93	*S. haemolyticus*	Oxacillin Gentamicin Mupirocin (I) Amoxicillin Cefoxitin Cefepime(I) Penicillin	Penam Aminoglycoside Mupirocin Penam Cephalosporin Cephalosporin Beta lactam	*mecA* *aph3-IIIa* *dfrG* *blaZ* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Aminoglycoside Diaminopyrimidine Penam Aminoglycoside
99	*S. haemolyticus*	Oxacillin Gentamicin Amoxicillin Tetracycline Cefoxitin Penicillin	Penam Aminoglycoside Penam Tetracycline Cephalosporin Penam	*mecA* *dfrG* *tetK* *lnu* *aac(6')-le-aph(2")-la* *blaZ* *aph3-IIIa*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Diaminopyrimidine Tetracycline Lincosamide Aminoglycoside Penam Aminoglycoside
105	*S. haemolyticus*	Oxacillin Gentamicin Amoxicillin Tetracycline Cefoxitin Cefepime Fusidic acid Penicillin	Penam Aminoglycoside Penam Tetracycline Cephalosporin Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *dfrG* *TetK* *Lnu* *aac(6')-le-aph(2")-la* *Aph3-IIIa*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Dianminopyrimidine Tetracycline Lincosamide Aminoglycoside Aminoglycoside

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
207	*S. hominis*	Oxacillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *msrA* *Ant(4)* *fusB* *qacA/B*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Macrolide Aminoglycoside Fusidic acid fluoroquinolone
208	*S. hominis*	Oxacillin Amoxicillin Erythromycin Tetracycline Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Tetracycline Cephalosporin Fusidic acid Penam	*mecA* *msrA* *mphC* *Aph3-IIIa* *Sat4A* *fusC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Streptogramin, Macrolide Macrolide Aminoglycoside Nucloside Fusidic acid
209	*S. hominis*	Oxacillin Amoxicillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *BlaZ* *msrA* *qacA/b* *ant(4')-lb* *fusB*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Macrolide Fluoroquinolone Aminoglycoside Fusidic acid
211	*S. cohnii*	Oxacillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *msrA* *mphC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Macrolide Macrolide
321	*S. epidermidis*	Oxacillin Amoxicillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *norA* *fusB* *mgrA* *qacA/B* *ermC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone, acridine dye Fusidic acid Peptide,penam, cephalosporin, acridinedye, fluoroquinolone, tetracycline Fluoroquinolone Lincosamide, streptogramin, macrolide

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
				dfrC *Ant(4')-lb*	Diaminopyrimidine Aminoglycoside
327	*S. epidermidis*	Oxacillin Gentamicin (I) Amoxicillin Cefoxitin Fusidic acid Penicillin	Penam Aminoglycoside Penam Cephalosporin Fusidic acid Penicillin	*mecA* *blaZ* *norA* *fusB* *mgrA* *qacA/B* *ermC* *dfrC* *Ant(4')-lb*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone, acridine dye Fusidic acid Peptide, penam, cephalosporin, acridine dye, fluoroquinolone, tetracycline Fluoroquinolone Lincosamide, streptogramin, macrolide Diaminopyrimidine Diaminopyrimidine Aminoglycoside
329	*S. epidermidis*	Oxacillin Amoxicillin Erythromycin Tetracycline Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Tetracycline Cephalosporin Fusidic acid Penam	*mecA* *BlaZ* *norA* *fusB* *mgrA* *qacA/B* *dfrC* *Ant(4')-lb*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone, acridine dye Fusidic acid Peptide, penam, cephalosporin, acridine dye, fluoroquinolone, tetracycline Fluoroquinolone Diaminopyrimidine Aminoglycoside
343	*S. cohnii*	Oxacillin Amoxicillin Erythromycin Tetracycline Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Tetracycline Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *mphC* *ermC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Macrolide Lincosamide, streptogramin, macrolide
349	*S. cohnii*	Oxacillin Amoxicillin Erythromycin Cefoxitin Fusidic acid	Penam Penam Macrolide Cephalosporin Fusidic acid	*mecA* *blaZ* *dfrC* *mphC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Diaminopyrimidine Macrolide

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
		Penicillin	Penam		
355	*S. epidermidis*	Oxacillin Amoxicillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *dfrC* *msrA* *mgrA* *norA* *qacA/B* *ant(4')-lb* *fusB*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Diaminopyrimidine Streptogramin, Macrolide Peptide, penam, cephalosporin, acridine dye, fluoroquinolone, tetracycline Fluoroquinolone, acridine dye Fluoroquinolone Aminoglycoside Fusidic acid
361	*S. haemolyticus*	Amoxicillin Cefoxitin Fusidic acid Penicillin	Penam Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *qacA/B* *fusB*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone Fusidic acid
372	*S. hominis*	Cefoxitin Penicillin	Cephalosporin Penam	*mecA* *blaZ*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam
373	*S. haemolyticus*	Oxacillin Amoxicillin Cefoxitin Penicillin	Penam Penam Cephalosporin Penam	*mecA* *blaZ* *dfrG* *fusB*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Diaminopyrimidine Fusidic acid
385	*S. hominis*	Cefoxitin Penicillin	Cephalosporin Penam	*mecA* *qacA/B*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Fluoroquinolone
386	*S. hominis*	Oxacillin Amoxicillin Erythromycin Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *qacA/B* *fusC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone Fusidic acid
387	*S. hominis*	Mupirocin Amoxicillin Erythromycin	Mupirocin Penam Macrolide	*mecA* *blaZ* *msrA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Streptogramin, Macrolide

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
		Cefoxitin	Cephalosporin	*mphC*	Macrolide
		Fusidic acid	Fusidic acid	*qacA/B*	Fluoroquinolone
		Penicillin	Penam	*fusC*	Fusidic acid
407	*S. epidermidis*	Oxacillin	Penam	*mecA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin
		Amoxicillin	Penam	*dfrC*	Diaminopyrimidine
		Cefoxitin	Cephalosporin	*norA*	Fluoroquinolone, acridine dye
		Fusidic acid	Fusidic acid	*mphC*	Macrolide
		Penicillin	Penam	*qacA/B*	Fluoroquinolone
				msrA	Streptogramin, Macrolide
435	*S. epidermidis*	Oxacillin	Penam	*mecA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin
		Gentamicin	Aminoglycoside	*blaZ*	penam
		Amoxicillin	Penam	*dfrC*	Diaminopyrimidine
		Cefoxitin	Cephalosporin	*ant(4')-lb*	Aminoglycoside
		Cefepime	Cephalosporin	*aac(6')-le-aph(2")-la*	Aminoglycoside
		Fusidic acid	Fusidic acid	*norA*	Fluoroquinolone, acridine dye
		Penicillin	Penam	*mupA*	mupirocin
				fusB	Fusidic acid
436	*S. epidermidis*	Oxacillin	Penam	*mecA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin
		Amoxicillin	Penam	*blaZ*	Penam
		Erythromycin	Macrolide	*tetK*	Tetracycline
		Tetracycline	Tetracycline	*norA*	Fluoroquinolone, acridine dye
		Cefoxitin	Cephalosporin	*ermC*	Lincosamide, streptogramin, macrolide
		Penicillin	Penam	*dfrC*	Diaminopyrimidine
				mgrA	Peptide, penam, cephalosporin, acridine dye, fluoroquinolone ,tetracycline
445	*S. haemolyticus*	Oxacillin	Penam	*mecA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin
		Mupirocin(I)	Mupirocin	*blaZ*	penam
		Amoxicillin	Penam	*msrA*	Streptogramin, Macrolide
		Erythromycin	Macrolide	*qacA/B*	Fluoroquinolone
		Cefoxitin	Cephalosporin	*cat*	Chloramphenicol
		Penicillin	Penam	*vgaA*	Streptogramin, pleuromutilin
		Chloramphenicol	Chloramphenicol	*mphC*	macrolide

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
465	*S. epidermidis*	Amoxicillin Erythromycin Tetracycline Cefoxitin Cefepime Fusidic acid Penicillin Chloramphenicol	Penam Macrolide Tetracycline Cephalosporin Cephalosporin Fusidic acid Penam Chloramphenicol	*mecA* *blaZ* *fusB* *norA* *dfrC* *tetK* *mphC* *msrA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fusidic acid Fluoroquinolone, acridine dye Diaminopyrimidine Tetracycline Macrolide Streptogramin, Macrolide
475	*S. epidermidis*	Oxacillin Mupirocin Erythromycin Cefoxitin Cefepime Fusidic acid Penicillin	Penam Mupirocin Macrolide Cephalosporin Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *norA* *dfrC* *mupA* *ermC*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone, acridine dye Diaminopyrimidine Muprcion Lincosamide, streptogramin, macrolide
479	*S. hominis*	Oxacillin Amoxicillin Erythromycin Tetracycline Cefoxitin Fusidic acid Penicillin	Penam Penam Macrolide Tetracycline Cephalosporin Fusidic acid Penicillin	*mecA* *blaZ* *ermC* *TetK* *fusB* *ant(4')-lb* *lnu* *qacA/B*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Lincosamide, streptogramin, macrolide Tetracycline Fusidic acid Aminoglycoside Lincosamide Fluoroquinolone
492	*S. haemolyticus*	Oxacillin Cefoxitin Penicillin	Penam Cephalosporin Penam	*mecA* *mphC* *msrA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin macrolide Streptogramin, macrolide
506	*S. haemolyticus*	Oxacillin Amoxicillin Erythromycin Tetracycline Cefoxitin	Penam Penam Macrolide Tetracycline Cephalosporin	*mecA* *blaZ* *ermA* *tetK* *qacA/B*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Lincosamide, streptogramin, macrolide Tetracycline Fluoroquinolone

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
		Cefepime Penicillin	Cephalosporin Penam		
538	*S. haemolyticus*	Oxacillin Amoxicillin Erythromycin(I) Cefoxitin Cefepime Fusidic acid Penicillin Chloramphenicol	Penam Penam Macrolide Cephalosporin Cephalosporin Fusidic acid Penicillin Chloramphenicol	*mecA* *blaZ* *fusB* *qacA/B*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fusidic acid Fluoroquinolone
620	*S. hominis*	Oxacillin Amoxicillin Cefoxitin Penicillin	Penam Penam Cephalosporin Penam	*mecA* *blaZ*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam
623	*S. hominis*	Oxacillin Amoxicillin Cefoxitin Penicillin	Penam Penam Cephalosporin Penam	*mecA* *blaZ*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam
631	*S. epidermidis*	Oxacillin Gentamicin Amoxicillin Erythromycin Cefoxitin Cefepime(I) Fusidic acid Penicillin	Penam Aminoglycoside Penam Macrolide Cephalosporin Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *ermA* *vgaB* *norA* *vatB* *dfrC* *qacA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Lincosamide, streptogramin, macrolide Streptogramin, pleuromutilin Fluoroquinolone, acridine dye Streptogramin Diaminopyrimidine Fluoroquinolone
664	*S. epidermidis*	Oxacillin Amoxicillin Cefoxitin Penicillin	Penam Penam Cephalosporin Penam	*mecA* *blaZ* *qacA* *ant(4')-lb* *norA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Fluoroquinolone Aminoglycoside Fluoroquinolone, acridine dye

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
				dfrC	Diaminopyrimidine
673	*S. epidermidis*	Oxacillin Gentamicin Amoxicillin Erythromycin Tetracycline Cefoxitin Fusidic acid Penicillin	Penam Aminoglycoside Penam Macrolide Tetracycline Cephalosporin Fusidic acid Penam	*mecA* *blaZ* *vgaB* *norA* *vatB* *aac(6')-le-aph-(2")-la* *ermA* *fusB* *qacA/B* *dfrC* *tetk*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Streptogramin, pleuromutilin Fluoroquinolone, acridine dye Streptogramin Aminoglycoside Lincosamide, streptogramin, macrolide Fusidic acid Fluoroquinolone Diaminopyrimidine Tetracycline
699	*S. warneri*	Oxacillin Amoxicillin Cefoxitin Penicillin	Penam Penam Cephalosporin Penam	*mecA* *blaZ* *msrA* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Streptogramin, macrolide Aminoglycoside
700	*S. warneri*	Oxacillin Amoxicillin Erythromycin Cefoxitin Penicillin	Penam Penam Macrolide Cephalosporin Penam	*mecA* *blaZ* *msrA* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Streptogramin, macrolide Aminoglycoside
702	*S. warneri*	Oxacillin Amoxicillin Erythromycin Cefoxitin Penicillin	Penam Penam Macrolide Cephalosporin Penam	*mecA* *blaZ* *msrA* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Streptogramin, macrolide Aminoglycoside
711	*S. epidermidis*	Oxacillin Gentamicin Amoxicillin Erythromycin	Penam Aminoglycoside Penam Macrolide	*mecA* *blaZ* *cat* *norA*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Chloramphenicol Fluoroquinolone, acridine dye

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
		Tetracycline Cefoxitin Penicillin Chloramphenicol	Tetracycline Cephalosporin Penam Chloramphenicol	*aac(6')-le-aph(2")-la* *mupA* *dfrC* *ermC* *qacA/B* *ant(4')-lb*	Aminoglycoside mupirocin Diaminopyrimidine Lincosamide, streptogramin, macrolide Fluoroquinolone Aminoglycoside
712	*S. epidermidis*	Oxacillin Gentamicin Amoxicillin Erythromycin Tetracycline Cefoxitin Penicillin Chloramphenicol	Penam Aminoglycoside Penam Macrolide Cephalosporin Penam Chloramphenicol Tetracycline	*mecA* *blaZ* *cat* *norA* *aac(6')-le-aph(2")-la* *mupA* *dfrC* *ermC* *qacA/B* *ant(4")-lb*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Chloramphenicol Fluoroquinolone, acridine dye Aminoglycoside mupirocin Diaminopyrimidine Lincosamide, streptogramin, macrolide Fluoroquinolone Aminoglycoside
713	*S. epidermidis*	Oxacillin Gentamicin Amoxicillin Erythromycin Tetracycline Cefoxitin Penicillin Chloramphenicol	Penam Aminoglycoside Penam Macrolide Cephalosporin Penam Chloramphenicol Tetracycline	*mecA* *blaZ* *cat* *norA* *aac(6')-le-aph(2")-la* *mupA* *dfrC* *ermC* *qacA/B* *ant(4")-lb*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Chloramphenicol Fluoroquinolone, acridine dye Aminoglycoside mupirocin Diaminopyrimidine Lincosamide, streptogramin, macrolide Fluoroquinolone Aminoglycoside
715	*S. epidermidis*	Oxacillin Mupirocin Amoxicillin Erythromycin Tetracycline	Penam Mupirocin Penam Macrolide Cephalosporin	*mecA* *blaZ* *cat* *norA* *aac(6')-le-aph(2")-la*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Chloramphenicol Fluoroquinolone, acridine dye Aminoglycoside

Isolate ID	Species	Phenotype	Drug group	Genotype	Drug group
		Cefoxitin Penicillin Chloramphenicol	Penam ChloramphenicoTetracycline	*mupA* *dfrC* *ermC* *qacA/B* *ant(4")-lb*	mupirocin Diaminopyrimidine Lincosamide, streptogramin, macrolide Fluoroquinolone Aminoglycoside
716	*S. epidermidis*	Oxacillin Mupirocin Amoxicillin Erythromycin Tetracycline Cefoxitin Penicillin Chloramphenicol	Penam Mupirocin Penam Macrolide Cephalosporin Penam Chloramphenicol Tetracycline	*mecA* *blaZ* *Tet(K)* *cat* *norA* *aac(6')-le-aph(2")-la* *mupA* *dfrC* *ermC* *qacA/B* *ant(4")-lb*	Penam, monobactam, cephalosporin, carbapenem, cephamycin Penam Tetracycline Chloramphenicol Fluoroquinolone, acridine dye Aminoglycoside mupirocin Diaminopyrimidine Lincosamide, streptogramin, macrolide Fluoroquinolone Aminoglycoside

Table 3.9: Summary of the resistance phenotypes and genotypes in staphylococci and the relevant antibiotic group they were resistant to

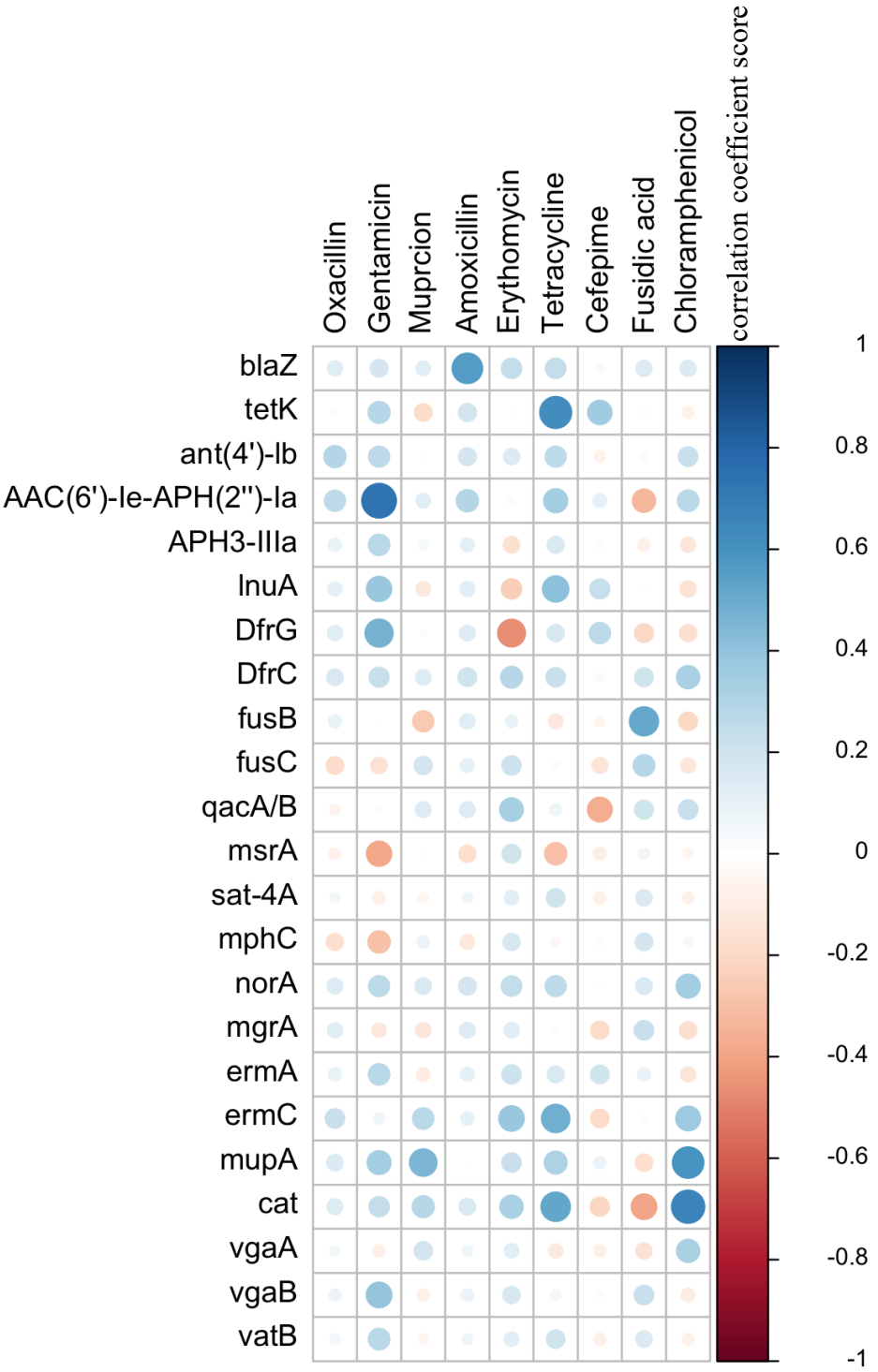


Figure 3.4: Parson correlation matrix of antibiotic resistance phenotype compared with antibiotic resistance genotype. White spaces not significantly correlated (<p=05). Blue circles indicated significant positive correlation and red show significant negative correlation. The size and strength of colour represent the numerical.

3.3.9 Determination of SCC*mec* types using WGS data

The SCC*mec* types were determined in 49 isolates that were *mecA*$^{+}$ by mapping for genetic markers from whole genome sequencing data (Table 3.10). 17 (34.7%) of 49 isolates which harboured the previously reported SCC*mec* types. These included SCC*mec* type IV (n=11) which was exclusively found in *S. epidermidis* isolates from public areas in hospitals; followed by type V (n=5) found in *S. haemolyticus* and *S warneri* and type VIII (n=1) found in an *S. hominis* isolate. The SCC*mec* element was absent in the genome of 10 (18.4%) isolates. 2 (4.1%) isolates harboured pseudo-SCC*mec* as they had *mec* complex but lacked the *ccr* complex. The remaining 19 (38.8%) isolates SCC*mec* types were untypable as they either have a novel combination of *mec* and *ccr* complex (n=4), or had multiple *ccr* complexes (n=13) or had novel *ccr* complexes (n=2). A select few of these SCC*mec* structures can be seen in Figure 3.5.

Isolate no	Area	Species	*mec* complex	*ccr* complex	SCC*mec* type
1	ELC	*S. haemolyticus*	C2	C	V
27	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
33	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
59	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
74	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
75	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
93	ELC	*S. haemolyticus*	Untypable	Absent	Pseudo
99	ELC	*S. haemolyticus*	C2	C A1/B1	Untypable
105	ELC	*S. haemolyticus*	C2	C	V
109	ELC	*S. sciuri*	Absent	Absent	No SCC*mec* element
207	WLC	*S. hominis*	A	Absent	Pseudo
208	WLC	*S. hominis*	A	A1/B1, A4/B4	Untypable
209	WLC	*S. hominis*	A	A1/B1, A4/B4	Untypable
211	WLC	*S. cohnii*	A	A1/B3	Untypable
321	ELH	*S. epidermidis*	B	A2/B2	IV
327	ELH	*S. epidermidis*	B	A2/B2	IV
329	ELH	*S. epidermidis*	B	A2/B2	IV
343	ELH	*S. cohnii*	A	A1, A3/B3	Untypable
349	ELH	*S. cohnii*	A	A1, A3/B3	Untypable
355	ELH	*S. epidermidis*	A	C, A2/B2	Untypable
361	ELH	*S. haemolyticus*	C2	Absent	Pseudo
372	ELH	*S. hominis*	A	A1/B1	Untypable
373	ELH	*S. haemolyticus*	Untypable	Absent	Pseudo
385	ELH	*S. hominis*	A	C, A1/B3	Untypable
386	ELH	*S. hominis*	A	A1/B1	Untypable
387	ELH	*S. hominis*	C2	A1/B1	Untypable
407	ELH	*S. epidermidis*	C2	C, A2/B2	Untypable
435	WLH	*S. epidermidis*	A	C, A2/B2	Untypable
436	WLH	*S. epidermidis*	B	C, A3/B3/, A4/B4	Untypable
445	WLH	*S. haemolyticus*	A	A2/B2	IV
465	WLH	*S. epidermidis*	B	A1/B1	Untypable

Isolate no	Area	Species	*mec* complex	*ccr* complex	SCC*mec* type
475	WLH	*S. epidermidis*	A	A2/B2	IV
479	WLH	*S. hominis*	A	A4/B4	VIII
492	WLH	*S. haemolyticus*	Untypable	C B4/A4	Untypable
506	WLH	*S. haemolyticus*	C2	Absent	Pseudo
538	WLH	*S. haemolyticus*	C2	A4/B4	Untypable
620	WLH	*S. hominis*	A	A1/B2	Untypable
623	WLH	*S. hominis*	B	A1/B2	Untypable
631	WLH	*S. epidermidis*	B	A2/B2	IV
664	WLH	*S. epidermidis*	B	C A2/B2	Untypable
673	WLH	*S. epidermidis*	C2	C A2/B2	Untypable
699	WLH	*S. warneri*	C2	C	V
700	WLH	*S warneri*	C2	C	V
702	WLH	*S warneri*	B	C	V
711	WLH	*S. epidermidis*	B	A2/B2	IV
712	WLH	*S. epidermidis*	B	A2/B2	IV
713	WLH	*S. epidermidis*	B	A2/B2	IV
715	WLH	*S. epidermidis*	B	A2/B2	IV
716	WLH	*S. epidermidis*	B	A2/B2	IV

Table 3.10: The diversity of SCC*mec* types of the 49 coagulase negative staphylococcal isolates recovered from public areas from the community and general public areas in hospitals.

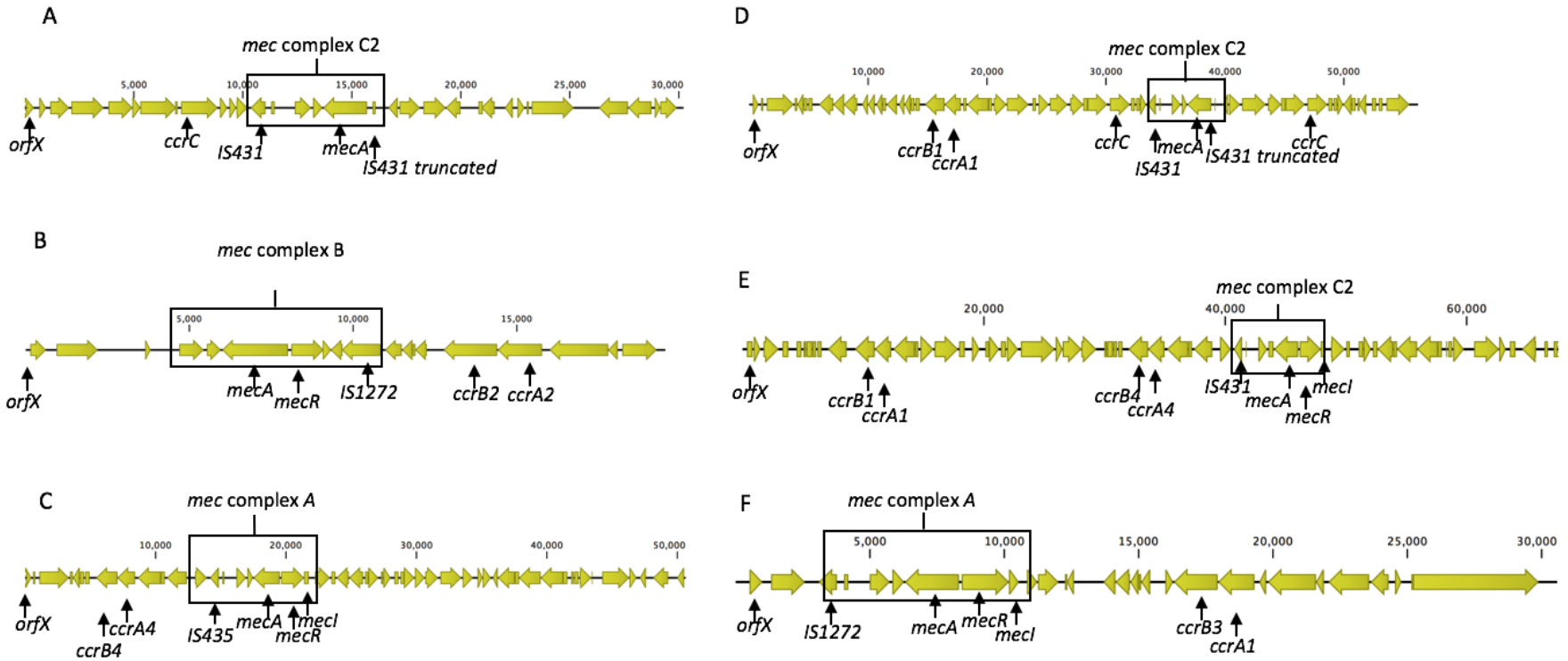


Figure 3.5: A select sample of SCC*mec* structure from staphylococcal isolates from high-frequency touched sites. A= Isolate 1: *S. haemolyticus* SCC*mec* type V; B= Isolate 475: *S. epidermidis* SCC*mec* type IV, C=479 S. *hominis* SCC*mec* type VIII; D=99 *S. haemolyticus* with *mec* C2 complex and *ccrC, ccrA1/B1* complex E= 208 *S. hominis* with a *ccrA1/B1, ccrB4/A4 complex* and F= 211 *S. cohnii* with a *mec* A complex and a *ccrB3/A1* complex. Figure in Cave et al. 2019

3.4 Discussion

3.4.1 Multidrug resistant staphylococcal species

281 multidrug resistant staphylococcal isolates belonging to 11 species were identified in this study. The most prevalent species were *S. epidermidis* (n=74) and *S. haemolyticus* (n=61). *S. epidermidis* and *S. haemolyticus* have previously been reported as the most common CoNS isolated from surfaces in public settings and hospitals surfaces (Seng et al., 2017b, 2017a; Xu et al., 2015). Interestingly, it has been demonstrated that *S. aureus (*n=5) was less prevalent on high-frequency hand touched surfaces, even though 30% of the human population are *S. aureus* carriers (Sollid et al., 2014). This may be because *S. aureus* is more commonly carried in the nasal passages than on hands (Tammelin et al., 2003). *S. aureus* is the most virulent species of staphylococci and the most common cause of infection in hospitalised patients (Liu, 2009). However, *S. epidermidis*, *S. hominis* and *S. haemolyticus* are amongst the most frequent nosocomial pathogens responsible for minor skin infections to life-threatening diseases (Basaglia et al., 2003; Huebner and Goldmann, 1999). In addition, community-associated CoNS have also been reported to cause infections (Chu et al., 2008).

3.3.2 MALDI-TOF MS validation

MALDI-TOF MS is the clinical standard for rapidly identifying bacteria to a species level (Schubert and Kostrzewa, 2017). Multiple companies which develop similar instruments, each building their preassembled database of microbial mass spectral profiles, however, the research conducting an experiment requires correct prediction at the species level to high confidence (Veloo et al., 2017). With few exceptions, there was excellent congruence between data derived using ASTA's Tinkerbell LT and Bruker's Autoflex at species level using the exact same isolates spotted on the target plate. When the two instruments mass spectrum were compared, the same key mass ions was observed even

in isolates which were identified as different species or their level of confidence in identification between the two instruments were different, suggesting the absence of a reference spectrum. These instruments are mainly used in clinical laboratories; therefore, their databased are assembled from clinically relevant isolates where environmental isolates were not considered a priority (Seuylemezian et al., 2018). However, the boundaries between isolates which are considered to be clinical or environmental isolates are becoming more blurred and gradually more environmental isolates will be incorporated in databases for taxa such as mycobacteria, staphylococci, streptococci (Clark et al., 2013).

3.4.3 Resistance phenotype

Amongst the staphylococcal isolates, there was an increased susceptibility toward penicillin (80.4%), fusidic acid (72.4%), and erythromycin (54.5%). Xu and co-workers reported increased susceptibilities toward penicillin, fusidic acid, erythromycin, and cefepime among staphylococcal isolates recovered from surfaces of inanimate objects in London hotel rooms (Xu et al., 2015). It has been reported that in primary care in England, 48.8% of antibiotics prescribed were penicillin and 13.4% were macrolides, lincosamides and streptogramins (Dolk et al., 2018). This potentially suggests why penicillin and erythromycin, a macrolide class antibiotic as two of the three most common antibiotic resistant phenotype from general public settings.

3.4.4 Comparison of the proportion of antibiotic resistant bacteria from East and West London

Areas in East and West London harboured high levels of antibiotic resistant staphylococci in proportion to the number of isolates that were examined. Significantly higher proportion ($P=0.0002$) of multidrug resistant staphylococci was observed from East London (56.7%) compared to West London (50.0%). This may be due to East London

having a higher population density ($9.7x10^3$ per km^2; 2017 estimate) compared to West London ($8.9x10^3$ per km^2; 2017 estimate) (Park, 2017). Previous studies have shown that there is a linkage in population density to the development of antibiotic resistant (Bruinsma et al., 2003).

There was no difference in distribution of these multidrug resistant isolates in two geographical areas at species level, apart from the observation that *S. warneri* isolates were exclusively recovered from West London, but not from East London, whereas *S. sciuri* and *S. equorum* were recovered from East London.

3.4.5 Comparison of the proportion of antibiotic resistant bacteria from public settings and public areas in hospitals

There was a high level of multidrug resistant staphylococci isolated in public areas in hospitals and general public settings. This was demonstrated by the number of isolates that were recovered. Statistically, there was a significantly higher proportion (P=0.0458) of multidrug resistant staphylococci in public areas in hospitals (49.5%) compared to that in general public settings (40.7%) which was expected due to the increased use of antibiotics in hospitals than in the community (Cantón and Morosini, 2011). However, the proportions of multidrug resistant bacteria isolated from general public settings in this study (46.8%) were less than that reported in similar studies from a university campus in Thailand (61%) and hotel rooms in London (86%) (Seng et al., 2017b; Xu et al., 2015). In this study, isolates were recovered from areas in hospitals that are accessible to the general public and not just to the hospital staff or patients. These areas included reception areas, public washrooms, corridors and lifts. The high levels of multidrug resistant staphylococci recovered from these areas in hospitals suggest a cross-contamination between community-associated and hospital-associated staphylococci.

3.4.6 Proportion of *mecA* gene by PCR from different areas

There was no significant difference in the carriage of the *mecA* gene in isolates recovered from East (10.71%) and West London (7.2%) and general public settings (7.7%) and public areas in a hospital (7.2%). The prevalence of the *mecA* gene in general public settings was less than that reported from the university campus in Thailand (20.5%) and hotel rooms in London (29.6%) (Seng et al., 2017b; Xu et al., 2015). In this study, the prevalence of the *mecA* gene in hospitals was also less than reported from a hospital in Thailand (70.1%). For the latter, it was expected as the isolates were recovered from medical ward surfaces and one would expect to find a high level of methicillin-resistance due to the frequent use of multiple antibiotics (Seng et al., 2017a).

Interestingly, 6 *S. sciuri* isolates that were *mecA*$^{+}$ by PCR and were resistant to oxacillin had a homolog of *mecA* known as *mecA1* (table 3.3). *mecA1* is considered to be the ancestry gene of *mecA*, which normally does not have resistance towards oxacillin. A recent study has shown that *S. sciuri* has developed oxacillin resistance using a variety of mechanisms from diversification of the non-binding domain of native PBPs, change in the *mecA* promoter, acquiring the SCC*mec* element and the adaptation of the bacterial genetic background (Clark et al., 2013)

3.4.7 Assembly statistics

The overall assembly of all the isolates that were WGS was too a high standard as they were similar in size and GC % to that of known reference isolates. The genome size and GC content of the assembly is similar to reference genomes *S. epidermidis* ATCC1228 (length 2,570,371bp, GC 32.1%), *S. haemolyticus* JCSC1435 (length 2,685,015 bp, GC 32.8%) *S. hominis* K1 (2,253,412, GC 31.4); *S. cohnii* FDAARGOS_334 (length 2,557,319, GC 32.7%) , *S. warneri* SG1 (length 2,486,042 GC 32.7%) and *S. sciuri* 285

(length 2,764,040; GC 32.6%) (Cheng et al., 2013; MacLea and Trachtenberg, 2017; "*Staphylococcus cohnii* (ID 24233) - Genome - NCBI," n.d.; "*Staphylococcus hominis subsp. hominis* (ID 2014) - Genome - NCBI," n.d.; "*Staphylococcus sciuri* (ID 10782) - Genome - NCBI," n.d.; Takeuchi et al., 2005)

3.4.8 Resistance genes

There was a large diversity of antibiotic resistance genes which encodes resistance to different types of antibiotics. Of these genes *blaZ* (90.7%) and *qacA/B* (51.2%) were the most common. Previous studies on the prevalence of antimicrobial resistance genes in CoNS from clinical and environmental is limited, but some reports have shown that *blaZ* is of one the most common antibiotic resistance genes found in staphylococci (Klibi et al., 2018; Pedroso et al., 2018). *QacA/B* has been previously reported to high prevalence from the University campus in Thailand (60.4%). This gene may have an important role for the survival of the bacteria within the environment as they encode multidrug efflux pump which has shown cross resistance-towards antiseptic and disinfectant compounds used to reduce bacterial contamination from surfaces (Wang et al., 2008).

Although *S. epidermidis* isolates were recovered from different areas, they possessed relatively similar antibiotic resistance profiles when compared by hierarchy clustering analysis. This may be due to the observation that all isolates had the fluoroquinolone efflux transporter gene *norA* and trimethoprim resistance dihydrofolate reductase gene *dfrC* (Costa et al., 2019; Totake et al., 1998). These genes may be essential for *S. epidermidis* survival, especially as *norA* like *qacA/B* has shown reduce susceptibility to antiseptic and disinfectant substances (Costa et al., 2019).

3.4.9 Correlation of antibiotic phenotype and genotype

For the isolates that had the *mecA* gene (WGS analysis), there were a few discrepancies with antibiotic phenotype to the predicted antibiotic resistance genotype. Seven antibiotic resistance genotypes correlated to the predicted genotype. Oxacillin, cefoxitin and penicillin correlation was not measured for the known resistance gene *mecA* due to all the isolates having the gene, therefore automatically have a strong correlation. It is possible that these resistance phenotypes that did not match with the genotype were derived from unknown SNP or genes which have not previously been described.

3.4.10 *SCCmec* classification

SCC*mec* was detected in 36 out of the 49 isolates that were whole genome sequenced; however, SCC*mec* types were assigned only to 17 isolates. The most common type was SCC*mec* type IV (n= 11), followed by SCC*mec* type V (n=5). These results are consistent with previously reported studies of clinical and from environmental isolates (Seng et al., 2017b). In this study, SCC*mec* type IV was exclusively found in *S. epidermidis* isolates. This is in keeping with others reporting a high association between SCC*mec* type IV and *S. epidermidis* (Chen et al., 2017). SCC*mec* type V was associated with *S. haemolyticus* and *S. warneri* isolates but is mainly reported to be associated with *S. haemolyticus* in clinical isolates (Zong et al., 2011). SCC*mec* type VIII was the only other typeable SCC*mec* from this study.

The remaining SCC*mec* types were untypeable as they harboured a novel *ccr* complex or multiple *ccr* complexes. Multiple *ccr* complexes have previously been described in clinical and community-associated isolates but currently, this is the first report of these SCC*mec* types determined in the isolates recovered from the general public environments (Chen et al., 2017). It has been reported that multiple ccr complexes have been shown to produce more stable *mecA* mRNA transcription compared to single elements as well as having a better cell wall integrity (Chen et al., 2017). This suggests that isolates with

multiple *ccr* complexes may have increased susceptibilities to oxacillin or cefoxitin, however, they do not always correlate with their phenotypic data. This adaptation may help the bacteria to survive more extended periods under persistent antibiotic pressure.

The remaining SCC*mec* types were untypeable as they harboured a novel *ccr* complex or multiple *ccr* complexes. Multiple *ccr* complexes have previously been described in clinical and community-associated isolates but currently, this is the first report of these SCC*mec* types determined in the isolates recovered from the general public environments (Chen et al., 2017). It has been reported that multiple ccr complexes have been shown to produce more stable *mecA* mRNA transcription compared to single elements as well as having a better cell wall integrity (Chen et al., 2017). This suggests that isolates with multiple *ccr* complexes may have increased susceptibilities to oxacillin or cefoxitin, however, they do not always correlate with their phenotypic data. This adaptation may help the bacteria to survive more extended periods under persistent antibiotic pressure.

3.5 Conclusion

In this chapter, multidrug resistant staphylococcal isolates were collected from high-frequency hand touched surface from public settings from East and West London and compared with each other. Their antibiotic resistance genes and SCC*mec* type were identified using WGS which have not been previously done on isolates from public settings. From these analyses, there were novel findings. These findings were:

1. General public areas and common public areas in hospitals in London can be reservoirs for MDR staphylococci. These MDR bacteria can be found at high levels on high-frequency touched surfaces.
2. Penicillin, erythromycin and fusidic acid the most common antibiotic staphylococci were resistant to in public settings.

1. There was a higher proportion of MDR staphylococci from East London compared to West London. This may be because East London has a larger population than West London.
2. Public areas in hospitals have a higher portion of MDR staphylococci compared to general public settings. This is due to there being a higher abundance of antibiotic used in hospitals compared with public settings in the community.
3. MALDI-TOF species identification for the majority of staphylococcal species was accurate except for few species that were not clinically related.
4. There was a strong correlation between resistance phenotype and known resistance genes in WGS isolates recovered in this study. Those that have shown a weak correlation had unknown genes or SNPs that confer resistance, which has not been previously described.
5. A diverse range of SCC*mec* types was determined from general public settings and public areas in hospitals of which many were untypeable due to having either a novel *ccr* or an extra *ccr* complex. These SCC*mec* structures have not been previously reported in isolates recovered from environmental surfaces in general public settings.

Overall, these findings show that these isolates have the potential to spread antibiotic resistant staphylococci to different people via general public settings and have the potential to cause infections which are untreatable with antibiotics that are currently available.

Chapter 4: Comparative genomics of *mecA* positive *S. epidermidis* isolates recovered from public settings using the One Health approach; determining horizontal gene transfer in these isolates

4.1 Introduction

*Staphylococcus epidermidi*s is the most common skin commensal and is the prominent cause of nosocomial and device-associated infections (Otto, 2009). In general coagulase-negative staphylococci (CoNS) lack virulence determinants, which generally are responsible for aggression. Nevertheless, they have factors which support adherence and colonisation. For *S. epidermidis*, one of its crucial virulence properties is the ability to attach to medical devices, and implants. This requires the bacteria to have the ability to produce biofilm (Trampuz and Zimmerli, 2005; Ziebuhr et al., 2006). For this to happen, the bacteria cells adhere to a surface, in which the bacterial cells then accumulate, forming a three-dimensional multi-layer, multi-cellular structure (Büttner et al., 2015). This step is critical for *S. epidermidis* pathogenesis as well as its internalisation and persistence in the host cells. Finally, the biofilm structure is dissembled (Büttner et al., 2015). From the dissembled biofilm, a single cell can disseminate through the bloodstream to colonise and form new biofilms at different sites in the body (Otto, 2008). This can lead to sepsis, meningitis and endocarditis (Becker et al., 2014). *S. epidermidis* is an important reservoir of mobile genetic elements (MGE) including antibiotic resistance and virulence genes (Conlan et al., 2012; Hung et al., 2015; Miragaia et al., 2009). Reports have shown that the arginine catabolic mobile element (ACME), which is important to *S. aureus* USA300 fitness, originated in *S. epidermidis* and has been horizontally transferred across (Planet et al., 2013). Extensive studies have been performed on the genetic lineages of *S. epidermidis* but little is known about the genetic lineages of isolates recovered from high-frequency touched surfaces in general public areas within hospitals and horizontal gene

transfer (HGT) of antibiotic resistance and virulence determents (Chaudhry and Patil, 2016; Conlan et al., 2012).

4.2 Method

Horizontal gene transfer can occur in *S. epidermidis* via bacteriophages, plasmid, transposons pathogenicity island and chromosomal cassettes. Multiple software has been developed to identify these genetic elements which can be split into two methods, parametric and phylogenetic (Malachowa and DeLeo, 2010; Ravenhall et al., 2015). Parametric methods search the genome for sections that are significantly different from that of the genome average, for example, the guanine-cytosine content or codon-usage (Ravenhall et al., 2015). The phylogenetic methods examine evolutionary histories of genes and identify conflicting phylogenies (Ravenhall et al., 2015). The benefits of the parametric methods are that it does not require a closely related isolate to be used for comparisons. Drawbacks to this method are that it relies on the host genome to be uniform in its genetic makeup without accounting for intragenomic variability, which can lead to overprediction (Ravenhall et al., 2015). Phylogenetic methods are better at characterising the HGT event as they can identify the donor species (Ravenhall et al., 2015). A drawback to this method is that there could be conflicts in the phylogenies due to events not considered by the mode; heavily relying on reliable reference species trees which are not always available, and the computational time to process and reconstruct many genes and species trees (Ravenhall et al., 2015). Therefore, in this study, multiple software (HGTector and Islandviewer 4) was used that can utilise both methods to detect HGT.

Whole genome sequencing phylogenetic studies and pangenome studies are used to infer relationship and evolution history between isolates from the same species (Caputo et al., 2019; McNally et al., 2016). WGS phylogenetic analyses can infer the evolutionary difference in bacterial genomes between isolates by looking for SNPs. Whereas in

pangenome analysis the genes from all isolates are clustered based on their genetic relativeness and then further grouped into core genes (found in all the isolates) or accessory genes (not always present in) within a bacterial species. Pangenome analysis can be used to characterise strains by a particular group of genes, for example, antibiotic resistance genes or virulent genes as well as determine the function of genes found in the core or accessory genome (Manara et al., 2018; Page et al., 2015). Additionally, the accessory genome distance can be interpreted in all isolates of a bacterial collection with a k-mer approach (Lees et al., 2019). This method is quick and can infer the difference in the accessory genome to a higher degree compared to pangenome due to the ability to use variable-length k-mers to accurately resolve genetic divergence (Lees et al., 2019).

In this study, bioinformatic analysis was used to determine Multilocus sequence types (MLST) and predict virulence, and antibiotic resistance genes in *mecA* positive (*mecA*$^{+}$) S. *epidermidis* isolates recovered from public settings from East and West London. It was also identified HGT events within the genome and the potential donor organisms of horizontally transferred antibiotic resistance and virulence genes. Additionally, it was investigated if these genes were transferred via plasmids or by phages or by other mobile genetic elements.

A core SNP maximum likelihood phylogenetic tree was constructed using *mecA*$^{+}$ *S. epidermidis* isolates from public settings from East and West London and other *S. epidermidis* isolates reference and draft genomes from the European Nucleotide Archive (ENA) database recovered from clinical samples (blood, cerebrospinal fluid, urine); healthy humans skin, nares and nasopharynx; livestock (cows, pigs and sheep); rodents (rats), plants, hospital environment from wards and catheters, animal housing and natural environment; to identify the genetic lineages of isolates from public settings. It was measured the distance of the accessory genome of isolates recovered from public setting

with the same isolates from the ENA database used in the core SNP maximum likelihood phylogenetic tree to determine if isolates recovered from the same source share the same accessory genes and the accessory genome using a k-mer approach (Poppunk) (Lees et al., 2019). A pangenome was also constructed to identify core and accessory genes in the genome. As part of the pangenome analyses, the absence and presence of antibiotic resistance and virulence genes across all *S. epidermidis* were identified as well as the Clusters of Orthologous Groups (COG) function family of the unique genes found in isolates collected from public settings in East and West London.

4.3 Results

4.3.1 MLST of WGS *Staphylococcus epidermidis*

The MLST sequence types (ST) of the *mecA*$^{+}$ *S. epidermidis* from public settings in East and West London sequence were determined from whole genome sequencing data. 10 different sequence types (ST) were assigned to 17 *S. epidermidis* isolates (Table 4.1). ST2 was the most common (n=5) sequence type, followed by ST66 (n=3) and ST87 (n=2). Two new sequence types were identified which have been assigned ST771 and ST779 which have been submitted to the MLST database.

S. epidermidis ID	Area	Sequence type (ST)
321	ELH	66
327	ELH	66
329	ELH	66
355	ELH	558
407	ELH	59
435	WLH	188
436	WLH	771
465	WLH	54
475	WLH	5
631	WLH	87
664	WLH	779
673	WLH	87
711	WLH	2
712	WLH	2
713	WLH	2
715	WLH	2
716	WLH	2

Table 4.1: MLST types of *S. epidermidis* isolates recovered from public settings in East and West London. East London Hospital= ELH; West London Hospital = WLH

4.3.2 Virulence genes identified in *mecA*$^+$ *S. epidermidis* recovered from public settings in East and West London

28 virulence genes were identified in 17 of the *mecA*$^+$ *S. epidermidis* isolates recovered from East and West London, of which 10 were identified in all isolates from public settings (Table 4.2). All isolates harboured the *nuc* gene (encoding thermonuclease an enzyme that can hydrolyse the host cell DNA and RNA), *sspA* gene (encoding serine V8 protease enzyme which is involved detaches bacterial cells from colonised sites); *sspB* gene (encoding cystine protease enzyme which breakdown elastin, fibronectin and kininogen), *lip* and *geh* genes (encoding lipase enzyme which breakdown fatty acid);

sdrG gene (encoding Ser- Asp rich fibrinogen-binding proteins which is involved in bacterial adhesion); *atl* (encoding autolysin involved in bacterial adhesion), *ebp* (encoding elastin binding protein involved in bacterial adhesion), *hlb* gene (encoding beta hemolysin toxin) and *capB* and *capC* gene (encoding polyglutamic acid capsule for immune invasion). *sdrH, sdrF, sdrG* and *sdrE* genes (encoding Ser-Asp rich fibrinogen-binding proteins which are involved in bacterial adhesion) were found in 76.2%, 64.7% (for *sdrF* and *sdrG*) and 11.8% respectively. *Ebh* gene (encoding cell wall-associated fibronectin-binding protein involved in bacterial adhesion) was found in 94.1% isolates. *IcaA*, *icaB*, *icaC*, i*caD* and *icaR* genes (Intercellular adhesion proteins involved in biofilm formation) were found in 47.1% of the isolates. *EsaA*, *essA*, *essB*, *essC*, *esxA* and *esaB* genes (encoding for the Type VII secretion system involved in bacterial survival and long-term persistence) were found in 23.6% (for *EsaA*, *essA*, *essB*, *essC)* and 5.9% (for *esaB*) of the isolates. Interestingly, 29.4% and 23.5% of isolates had the *gtaB* gene (encoding polysaccharide capsule) and *cylR2* (encoding cytolysin which is involved in lysing erythrocytes, polymorphonuclear leukocytes and macrophages) respectively which are virulent factors normally associated with *Bacillus* and *Enterococcus* respectively.

Virulence gene	Function	Percentage (%) of the *mecA* isolates from public areas in East and West London
nuc	Thermonuclease	100
sspA	Serine V8 protease enzyme	100
lip	Lipase enzyme	100
geh	Lipase enzyme	100
sdrG	Ser- Asp rich fibrinogen-binding proteins	100
atl	Autolysin	100
ebp	Elastin binding protein	100
hlb	Beta hemolysin toxin	100
capB	Polyglutamic acid capsule	100
capC	Polyglutamic acid capsule	100
sdrH	Ser-Asp rich fibrinogen-binding proteins	11.8
sdrF	Ser-Asp rich fibrinogen-binding proteins	76.2
sdrG	Ser-Asp rich fibrinogen-binding proteins	64.7
sdrE	Ser-Asp rich fibrinogen-binding proteins	11.8
Ebh	Cell wall-associated fibronectin-binding protein	94.1
IcaA	Intercellular adhesion proteins	47.1
icaB	Intercellular adhesion proteins	47.1
icaC	Intercellular adhesion proteins	47.1
i*caD*	Intercellular adhesion proteins	47.1
icaR	Intercellular adhesion proteins	47.1
esaA	Type VII secretion system	23.6
essA	Type VII secretion system	23.6
essB	Type VII secretion system	23.6
essC	Type VII secretion system	23.6
esaB	Type VII secretion system	5.9
gtaB	polysaccharide capsule	29.4
cylR2	Cytolysin	23.5

Table 4.2: Percentage of Virulence genes found in the 17 *mecA*$^{+}$ *S. epidermidis* recovered from East and West London

A hierarchy clustering heatmap was used to group the virulence genes if they were isolated from public areas in hospitals in East or West London (Figure 4.1). Interestingly, isolates that had the *icaA, icaB, icaC, icaD* and *icaR* were from West London hospital where the *esaA*, *essA*, *essB*, *essC* and *esxA* were found in East London hospital isolates.

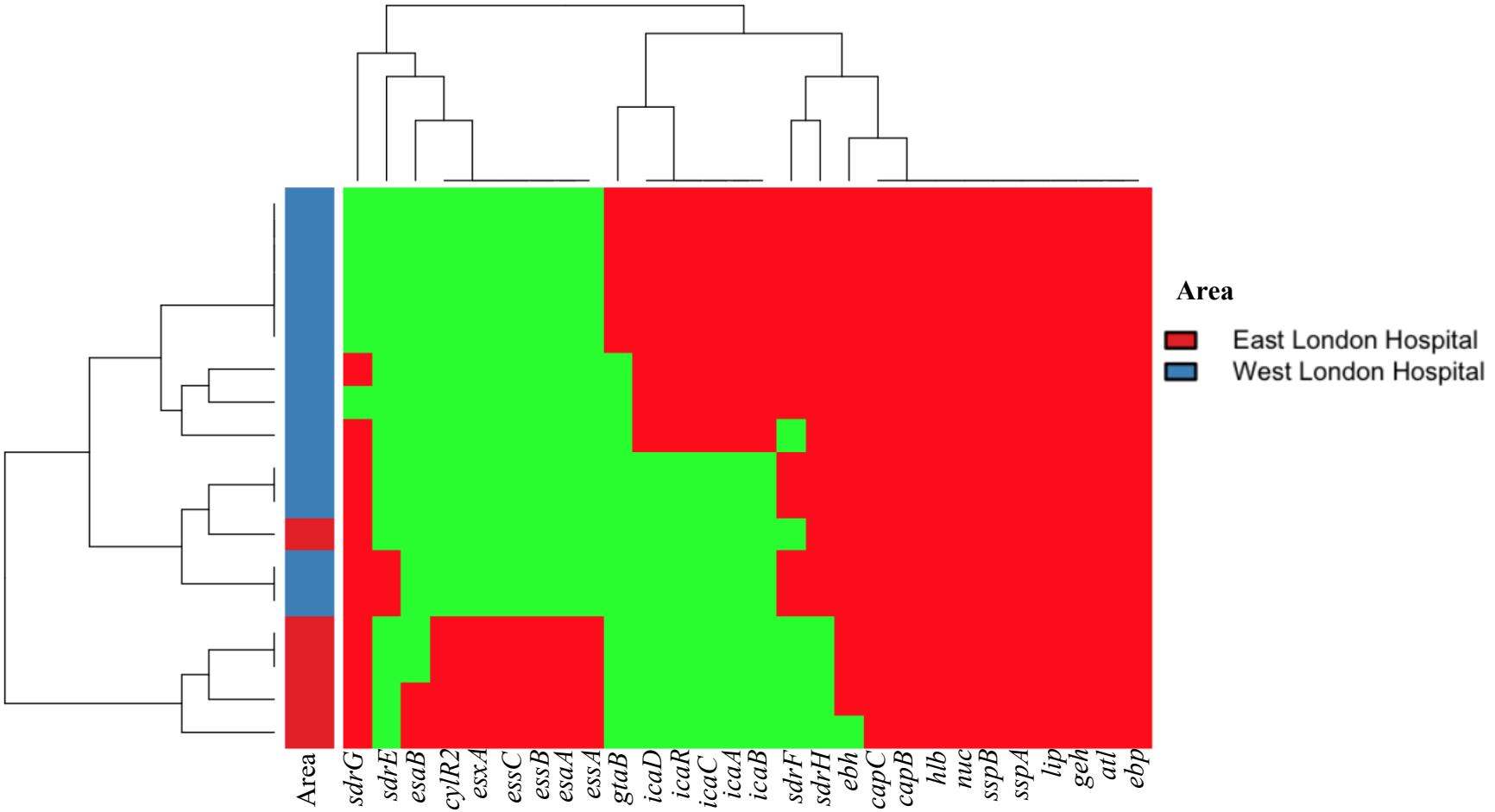


Figure 4.1: Hierarchy clustering heatmap of isolates virulence gene profiles in comparison to the area they were recovered from. (A)Hierarchy heatmap red tile present of virulent genes; green absent of virulent genes.

4.3.3 Horizontal gene transfer in isolates recovered from public settings in East and West London

Horizontally transferred genes were determined in 17 *mecA*$^{+}$ *S. epidermidis* isolates from public setting using the HGTector pipeline. 8.6% to 10.1% of the *S. epidermidis* isolates from public setting genes are considered to be horizontally transferred based on BLAST hit distribution patterns from NCBI non-redundant protein sequences database (Pruitt et al., 2007; Zhu et al., 2014) (Table 4.3). In comparison, reference isolate *S. epidermidis* ATCC 1228 recovered from human skin and mucosa was shown to have 8.6% of its gene to be horizontally transferred. 118 genera were predicted to be donors of HGT from the 17 *mecA*$^{+}$ *S. epidermidis* isolate recovered from public settings in London based on best hit of the non-redundant protein sequences database and NCBI taxonomy database of which *Bacillus* (mean n=34), *Macrococcus* (mean n=20) and *Salinicoccus* (mean n=19) genera had the most genes transferred from (table 4.4).

S. epidermidis Isolate No.	Number of protein-coding genes	Number of predicted HGT-derived genes	Percentage (%) of HGT derived genes
ATCC 12228	2416	208	8.6
321	2420	233	9.6
327	2419	231	9.6
329	2417	228	9.5
355	2308	208	9.0
407	2257	176	7.8
435	2453	248	10.1
436	2264	200	8.8
465	2267	194	8.6
475	2276	209	9.2
631	2250	216	9.6
664	2330	230	9.9
673	2276	208	9.1
711	2423	231	9.5
712	2422	230	9.5
713	2421	231	9.5
715	2388	219	9.2
716	2417	230	9.5

Table 4.3: Number of HGT genes predicted from the HGTector pipeline in *mecA*$^{+}$ *S. epidermidis* from East and West London and reference genome *S. epidermidis* ATCC 122

	***S. epidermidis* ID**																	
Predicted donor genus	**ATCC 11228**	**321**	**327**	**329**	**355**	**407**	**435**	**436**	**465**	**475**	**631**	**664**	**673**	**711**	**712**	**713**	**715**	**716**
Acidithrix	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
Actinomyces	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Aerococcus	1	1	2	2	2	0	2	1	1	1	3	0	2	1	1	1	1	0
Aeromicrobium	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Aeromonas	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amphibacillus	0	0	0	0	3	0	0	0	0	2	0	0	0	0	0	0	0	0
Amycolatopsis	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Andreesenia	1	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
Andreprevotia	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Aneurinibacillus	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0
Anthococcus	1	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	1	1
Aquibacillus	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Arachidicoccus	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Arcobacter	0	1	1	1	1	0	1	0	0	0	0	1	0	1	1	1	1	1
Atopostipes	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
Auricoccucs	4	7	7	6	2	2	5	3	6	5	6	6	6	6	5	5	6	5
Bacillus	29	33	34	33	29	27	37	29	32	31	30	29	32	40	39	37	40	39
Beduini	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
Bhargavaea	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Blautia	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Brevibacillus	0	1	1	1	1	0	0	2	2	2	0	0	0	0	0	0	0	0
Caenibacillus	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
Carnobacterium	2	4	3	3	4	2	4	6	2	2	2	2	2	3	3	2	2	2
Caryophanon	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Clostridium	4	3	3	3	4	3	6	3	5	4	4	5	5	3	3	3	3	3

	***S. epidermidis* ID**																	
Predicted donor genus	**ATCC 11228**	**321**	**327**	**329**	**355**	**407**	**435**	**436**	**465**	**475**	**631**	**664**	**673**	**711**	**712**	**713**	**715**	**716**
Collinsella	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0
Coprobacillus	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1
Corynebacterium	0	3	3	3	1	1	1	1	2	1	1	2	2	0	0	0	0	0
Curtobacterium	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Cutibacterium	0	1	1	1	2	2	1	0	2	1	1	3	1	3	3	3	3	3
Deferribacter	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1
Desmospora	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Desulfosporosinus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Desulfotomaculum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
Dickeya	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Domibacillus	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0
Drancourtella	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
Edaphobacillus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Eikenella	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Enterobacter	1	7	7	7	6	1	2	2	1	2	3	1	3	2	2	2	2	2
Enterococcus	16	10	10	10	10	2	11	11	6	11	13	7	13	9	8	8	8	8
Eremococcus	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Eubacterium	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1	1	1	1
Exiguobacterium	0	0	0	0	1	2	1	1	4	0	1	1	0	1	1	1	1	1
Fictibacillus	2	0	0	0	0	2	2	1	2	2	1	2	1	2	2	2	2	2
Gallibacterium	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	0	1
Gemella	0	0	0	0	0	0	2	0	0	0	0	0	0	2	2	2	2	2
Geobacillus	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0
Gracilibacillus	5	1	1	1	0	1	3	1	0	1	1	1	2	1	1	1	2	1
Halalkalibacillus	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Halanaerobium	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Halobacillus	0	0	0	0	1	0	1	0	0	0	0	0	0	1	1	1	1	1

	***S. epidermidis* ID**																	
Predicted donor genus	**ATCC 11228**	**321**	**327**	**329**	**355**	**407**	**435**	**436**	**465**	**475**	**631**	**664**	**673**	**711**	**712**	**713**	**715**	**716**
Halolactibacillus	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Helcococcus	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Jeotgalibacillus	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Jeotgalicoccus	8	10	9	10	14	7	8	9	2	9	6	11	6	6	7	7	6	6
Kurthia	0	0	0	0	0	1	1	0	0	0	0	1	0	1	1	1	1	1
Kyrpidia	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Lachnoanaerobaculum	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
Lachnoclostridium	0	0	0	0	0	1	1	0	1	0	0	0	0	1	1	1	1	1
Lactobacillus	8	11	11	11	11	8	10	9	8	10	8	10	7	9	10	9	8	10
Lactococcus	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1
Lentibacillus	0	2	2	2	0	0	1	3	3	0	0	0	0	0	0	0	0	1
Leptotrichia	0	2	2	2	1	0	0	2	0	0	0	0	0	0	0	0	0	0
Listeria	2	1	1	1	3	1	3	2	1	0	0	2	2	1	1	2	2	1
Lysinibacillus	2	6	6	6	3	2	3	2	4	2	2	3	2	1	1	2	2	2
Macrococcus	19	23	23	23	19	17	21	21	16	18	16	18	16	20	20	20	20	20
Mannheimia	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0
Marinilactibacillus	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Marinococcus	0	0	0	0	0	0	0	0	1	2	1	0	1	1	1	1	0	0
Massilibacterium	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0
Massilioclostridium	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
Methanobrevibacter	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Methylocaldum	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Microvirga	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1
Moraxella	0	0	0	0	0	1	1	0	1	0	0	0	0	1	1	1	1	1
Neisseria	0	1	1	0	0	0	0	1	0	0	0	0	0	1	1	1	0	1
Nocardia	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Nosocomiicoccus	4	5	5	4	2	2	8	4	4	4	5	3	6	4	3	4	2	4

	S. epidermidis ID																	
Predicted donor genus	**ATCC 11228**	**321**	**327**	**329**	**355**	**407**	**435**	**436**	**465**	**475**	**631**	**664**	**673**	**711**	**712**	**713**	**715**	**716**
Novibacillus	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0
Numidum	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1
Oceanobacillus	4	3	3	3	3	2	8	6	5	5	8	4	7	5	5	5	4	5
Oenococcus	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
Oleispira	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Ornithinibacillus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Paenibacillus	12	9	9	9	8	11	11	8	10	11	11	14	11	11	12	10	10	10
Parageobacillus	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Parvimonas	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	1	1	1
Planococcus	5	4	4	4	4	5	3	4	4	5	4	5	5	4	4	4	4	5
Planomicrobium	0	0	0	1	0	0	1	0	1	0	1	0	0	1	1	1	1	0
Pontibacillus	2	2	2	2	0	0	2	0	2	2	3	2	2	2	2	2	2	2
Pseudomonas	5	8	8	8	5	3	3	6	2	5	5	2	5	3	4	4	3	1
Psychrobacter	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0
Rhizobium	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Rodentibacter	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Ruminococcus	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Rummeliibacillus	1	0	0	0	0	1	1	1	1	2	2	0	2	1	1	1	1	1
Saccharibacillus	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Salinicoccus	18	24	24	23	15	20	21	15	13	19	17	20	15	20	19	21	18	18
Salipaludibacillus	1	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1
Sedimentibacter	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Segetibacter	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0
Sporolactobacillus	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Sporosarcina	0	1	1	1	2	2	2	0	2	0	2	0	1	0	0	0	0	0
Streptococcus	9	14	14	12	11	11	16	11	10	10	9	15	10	21	21	21	19	22
Streptomyces	1	1	0	1	1	1	3	3	2	3	2	3	4	2	1	2	1	4

	S. epidermidis ID																	
Predicted donor genus	**ATCC 11228**	**321**	**327**	**329**	**355**	**407**	**435**	**436**	**465**	**475**	**631**	**664**	**673**	**711**	**712**	**713**	**715**	**716**
Terribacillus	1	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	1	1
Tetragenococcus	1	2	2	2	3	1	4	1	2	1	3	1	3	1	1	1	1	1
Thalassobacillus	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Thermoactinomyces	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thiorhodovibrio	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichococcus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tuberibacillus	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1
Tumebacillus	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
Vagococcus	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
Veillonella	8	2	2	2	4	6	3	1	3	1	1	5	2	3	2	2	4	5
Virgibacillus	2	2	2	2	2	2	2	3	3	3	2	2	3	5	4	5	4	5
Weissella	1	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2
Youngiibacter	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 4.4: *MecA*$^{+}$ *S. epidermidis* and reference ATCC 11228 HGT-derived from by the best match putative donor genus as indicated by the best distal match.

13 out of 17 antibiotic resistance genes from the 17 *mecA*$^{+}$ *S. epidermidis* isolates were predicted to be horizontally transferred from another bacteria (Table 4.5). The *mecA* gene was predicted in all isolates to be horizontally transferred. The donor species was predicted as *Macrococcus canis.* Interestingly, *ermC, AAC(6')-Ie-APH(2'')-Ia,* and *mupA* are the only horizontally transferred genes that were predicted to be donated from two different species from different organisms. *ErmC* was found to be transferred from *Neisseria meningitidis MC58* except for isolate 475 in which it was predicted to be donated from *Actinomyces* spp. *S6-Spd3*. *AAC(6')-Ie-APH(2'')-Ia* was predicted to be transferred from *Enterococcus faecalis* V583 except for isolates 631 and 673 which was donated from *Streptococcus mitis B6. MupA* was predicted to be donated by *Bacillus halmapalus* except for isolate 435, 475 and 715 which was donated by Bacillus spp. *V-88*.

***S. epidermidis* ID.**	**Antibiotic resistance gene**	**Closest predicted donor species**
321	*mecA* *ermC* *ANT(4')-IB* *qacA/B* *fusB*	*Macrococcus canis* *Neisseria meningitidis MC58* *Arcobacter thereius* *Bacillus ndiopicus* *Enterococcus thailandicus*
327	*mecA* *qacA/B* *ermC* *ANT(4')-IB* *fusB*	*Macrococcus canis* *Bacillus ndiopicus* *Neisseria meningitidis MC58* *Arcobacter thereius* *Enterococcus thailandicus*
329	*mecA* *ermC* *ANT(4')-IB* *qacA/B* *fusB*	*Macrococcus canis* *Neisseria meningitidis MC58* *Arcobacter thereius* *Bacillus ndiopicus* *Enterococcus thailandicus*
355	*mecA* *ANT(4')-IB* *msrA* *qacA/B* *fusB*	*Macrococcus canis* *Arcobacter thereius* *Veillonella atypica* *Bacillus ndiopicus* *Enterococcus thailandicus*
407	*mecA* *mphC* *msrA* *qacA/B*	*Macrococcus canis* *Veillonella atypica* *Veillonella atypica* *Bacillus ndiopicus*

S. epidermidis ID.	Antibiotic resistance gene	Closest predicted donor species
435	*mecA*	*Macrococcus canis*
	mupA	*Bacillus* spp. *V-88*
	ant(4')-IB	*Arcobacter thereius*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis V583*
	fusB	*Enterococcus thailandicus*
436	*mecA*	*Macrococcus canis*
	tet(k)	*Streptomyces cinnamoneus*
	ermC	*Neisseria meningitidis MC58*
465	*mecA*	*Macrococcus canis*
	mphC	*Veillonella atypica*
	mrsA	*Veillonella atypica*
	tet(k)	*Lactobacillus kimchicus*
	fusB	*Enterococcus thailandicus*
475	*mecA*	*Macrococcus canis*
	ErmC	*Actinomyces* spp. *S6-Spd3*
	mupA	*Bacillus* spp. *V-88*
631	*vgaB*	*Bacillus tuaregi*
		Enterococcus spp.
	ermA	*HMSC29A04*
	mecA	*Macrococcus canis*
	qacA/B	*Bacillus ndiopicus*
	fusB	*Enterococcus thailandicus*
	vatB	*Bacillus gottheilii*
	AAC(6')-Ie-APH(2'')-Ia	*Streptococcus mitis B6*
664	*mecA*	*Macrococcus canis*
	ANT(4')-IB	*Arcobacter thereius*
	qacA/B	*Bacillus ndiopicus*
673	*vgaB*	*Bacillus tuaregi*
	mecA	*Macrococcus canis*
	qacA/B	*Bacillus ndiopicus*
	tet(k)	*Streptomyces cinnamoneus*
		Enterococcus spp.
	ermA	*HMSC29A04*
	fusB	*Enterococcus thailandicus*
	vatB	*Bacillus gottheilii*
	AAC(6')-Ie-APH(2'')-Ia	*Streptococcus mitis B6*
711	*mupA*	*Bacillus halmapalus*
	mecA	*Macrococcus canis*
	ANT(4')-IB	*Arcobacter thereius*
	ermC	*Neisseria meningitidis MC58*
		Streptococcus hyovaginalis
	cat	*DSM 12219*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis V583*
712	*mupA*	*Bacillus halmapalus*
	mecA	*Macrococcus canis*
	ANT(4')-IB	*Arcobacter thereius*
	ermC	*Neisseria meningitidis MC58*

S. epidermidis ID.	Antibiotic resistance gene	Closest predicted donor species
	cat	*Streptococcus hyovaginalis DSM 12219*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis*
713	*mupA*	*Bacillus halmapalus*
	mecA	*Macrococcus canis*
	ermC	*Neisseria meningitidis MC58*
	ANT(4')-IB	*Arcobacter thereius*
	cat	*Streptococcus hyovaginalis DSM 12219*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis*
715	*ANT(4')-IB*	*Arcobacter thereius*
	mupA	*Bacillus* spp. *V-88*
	mecA	*Macrococcus canis*
	cat	*Streptococcus hyovaginalis DSM 12219*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis*
716	*ANT(4')-IB*	*Arcobacter thereius*
	ErmC	*Neisseria meningitidis MC58*
	mecA	*Macrococcus canis*
	mupA	*Bacillus halmapalus*
	cat	*Streptococcus hyovaginalis DSM 12219*
	AAC(6')-Ie-APH(2'')-Ia	*Enterococcus faecalis*

Table 4.5: HGT antibiotic resistance genes and the predicted donor organism from *mecA*$^+$ *S. epidermidis* isolates from public areas from East and West London.

6 out of 28 virulence genes were predicted to be horizontally transferred (Table 4.6). The *hlb* which was ubiquitous in all the *mecA*$^+$ *S. epidermidis* isolates from East and West London was predicted to be horizontally transferred from *Thermoactinomyces* spp. In isolates 321, 327 and 329 the predicted donor of *hlb* gene was *Thermoactinomyces* spp. wherein the other isolates the predicted donor was *Virgibacillus alimentarius*. Interestingly, it was predicted the *icaA* and *icaR* genes which encode for the intercellular adhesion and the *essC* and the *esxA* gene which encodes for type VII secretion system to be horizontal transferred. The *icaR* gene was predicted to be donated by 3 species. In isolates 435, 711, 713 and 716 *icaR* were predicted to be donated from *Lachnospiraceae bacterium C6A11*; isolates 436 and 465 were predicted to be donated from *Lentibacillus*

jeotgali and in isolates 712 and 715 were predicted to be donated from *Desulfotomaculum acetoxidans* DSM 771.

***S. epidermidis* ID**	**gene**	**Closest predicted donor species**
321	*essC*	*Listeria booriae*
	esxA	*Bacillus* spp. *UNC41MFS5*
	hlb	*Thermoactinomyces* spp.
	cylR2	*Clostridium* spp. *W14A*
327	*essC*	*Listeria booriae*
	esxA	*Bacillus* spp. *UNC41MFS5*
	hlb	*Thermoactinomyces* spp.
	cylR2	*Clostridium* spp.
329	*essC*	*Listeria booriae*
	esxA	*Bacillus* spp. *UNC41MFS5*
	hlb	*Thermoactinomyces* spp.
	cylR2	*Clostridium* spp.
355	*essC*	*Listeria booriae*
	esxA	*Bacillus* spp. *UNC41MFS5*
	hlb	*Virgibacillus alimentarius*
	cylR2	*Weissella hellenica*
407	*hlb*	*Virgibacillus alimentarius*
435	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Lachnospiraceae bacterium C6A11*
	hlb	*Virgibacillus alimentarius*
436	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Lentibacillus jeotgali*
	hlb	*Virgibacillus alimentarius*
465	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Lentibacillus jeotgali*
	hlb	*Virgibacillus alimentarius*
475	*hlb*	*Virgibacillus alimentarius*
631	*hlb*	*Virgibacillus alimentarius*
664	*hlb*	*Virgibacillus alimentarius*
673	*hlb*	*Virgibacillus alimentarius*
711	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Lachnospiraceae bacterium C6A11*
	hlb	*Virgibacillus alimentarius*
712	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Desulfotomaculum acetoxidans DSM 771*
	hlb	*Virgibacillus alimentarius*
713	*icaA*	*Macrococcus caseolyticus JCSC5402*
	icaR	*Lachnospiraceae bacterium C6A11*
	hlb	*Virgibacillus alimentarius*
715	*icaA*	*Macrococcus caseolyticus*
	icaR	*Desulfotomaculum acetoxidans DSM 771*
	hlb	*Virgibacillus alimentarius*

716	*icaA* *icaR* *hlb*	*Macrococcus caseolyticus JCSC5402* *Lachnospiraceae bacterium C6A11* *Virgibacillus alimentarius*

Table 4.6: *MecA*$^{+}$ *S. epidermidis* isolates from public areas from East and West London HGT virulence genes and the predicted donor organism.

4.3.4 Antibiotic resistance and virulence genes carried within a genomic island

All *mecA*$^{+}$ *S. epidermidis* isolates recovered from East and West London were predicted to have multiple genomic islands based on Islandviewer 4 methods (Bertelli et al., 2017) (Table 4.7). Interestingly, the *mecA* gene was predicted to be found on a genomic island in 4 of the isolates. For the other isolates, the *mecA* gene was near a genomic island. Antibiotic resistance genes were detected in all isolates except for isolate 407. From these 17 *mecA*$^{+}$ *S. epidermidis* isolates 2 (11.8%) had 7 antibiotic resistance genes; 1 (5.9%) had 6 antibiotic resistance genes; 4 (23.53%) had 5 antibiotic resistance genes; 1 (5.9%) had 4 antibiotic resistance gene; 2 (11.8%) had 3 antibiotic resistance gene; 2 (11.8%) had 4 antibiotic resistance genes, and 1 (5.9%) had 1 antibiotic resistance gene which was predicted to be encoded within a genomic island regions. The most common antibiotic resistance gene detected within genomic island regions were *blaZ* (n=10); *qacA/B* and *AAC(6')-Ie-APH(2")-Ia* (n=7); *ANT(4")-Ib* (n=6); *ermC* (n=5); *mupA* (n=4); *vatB*, *vgaA*, *ANT(4")-Ib* (n=2) and then *fusB* and *tet(K)* (n=1). The *cylR2* gene in isolate 321, 327 and 329 were all found to be within a genomic island region.

S. epidermidis ID	No. of Predicted genomic island regions	Antibiotic genes	Virulence genes
321	8	*qacA/B* *ANT(4")-Ib*	*cylR2*
327	11	*ermC* *qacA/B*	*cylR2*
329	10	*qacA/B*	*cylR2*
355	5	*fusB* *qacAB* *ANT(4")-Ib*	*cylR2*
407	4		
435	7	*blaZ*	
436		*ErmC* *tetK* *blaZ*	
465	5	*blaZ*	
475	4	*blaZ* *mecA* *mupA* *blaZ*	
631	6	*AAC(6')-Ie-APH(2")-Ia* *vatB* *vgaB* *qacA/B* *ANT(4')-Ib* *ermA* *blaZ*	
664	5	*qacA/B* *ANT(4')-Ib*	
673	7	*AAC(6')-Ie-APH(2")-Ia* *vgaB* *ermA* *vatB* *tet(k)* *qacA/B* *blaZ*	
711	10	*cat* *AAC(6')-Ie-APH(2")-Ia* *blaZ* *mupA* *ermC*	
712	8	*mecA* *AAC(6')-Ie-APH(2")-Ia* *cat* *ermC* *mupA*	

S. epidermidis ID	No. of Predicted genomic island regions	Antibiotic genes	Virulence genes
		blaZ	
713	7	*cat* *blaZ* *ermC* *AAC(6')-Ie-APH(2")-Ia* *mecA*	
715	7	*AAC(6')-Ie-APH(2")-Ia* *mupA* *ANT(4')-Ib* *mecA* *blaZ*	
716	8	*cat* *AAC(6')-Ie-APH(2")-Ia* *blaZ* *ermC* *ANT(4')-Ib*	

Table 4.7: Genomic islands of 17 *mecA*+ *S. epidermidis* isolates from East and West London

4.3.5 Antibiotic and virulence resistance genes carried on plasmids

All *mecA*$^{+}$ *S. epidermidis* (n=17) isolates recovered from East and West London possessed plasmids (ranging from 1 to 4) (Table 4.8). 15 of the 17 isolates carried one resistance gene. Three isolates (17.7%) had 5 antibiotic resistance genes carried on their plasmids; 3 isolates (17.7%) had 4 antibiotic resistance gene; 6 isolate (35.3%) had 3 antibiotic resistance genes, 2 isolates had 2 antibiotic resistance gene, 1 isolate had 1 antibiotic resistance gene and 2 isolates had 2 antibiotic resistance genes carried on their plasmids. Two isolates had more antibiotic resistance genes encoded on their plasmids than their chromosome (isolates 321 and 435). Up to 4 antibiotic resistance genes were carried on a single plasmid. *ErmC* and *ant(4')-Ib* (n=10) was the most common resistance genes to be carried on a plasmid followed by *qacA/B*, *cat*, and *fusB* (n=5); then *blaZ* (n=4); *mupA* (n=3); *AAC(6')-le—APH(2")-la, mphC* and *msrA* (n=1). No virulence genes were identified to be carried on the plasmids. Interestingly, *blaZ* was not predicted to be

donated from a different organism but was identified to be encoded on some of the isolate's plasmids.

***S. epidermidis* ID**	**Plasmids**	**Chromosome: plasmid-encoded antibiotic genes**	**Antibiotic resistance genes (and plasmid they are found on)**
321	3	4:5	*ermC* (plasmid 1) *blaZ* (plasmid 2) *fusB* (plasmid 2) *ant(4')-Ib* (plasmid 2) *qacA/B* (plasmid 2)
327	2	5:4	*blaZ* (plasmid 1) *fusB* (plasmid 1) *ant(4')-Ib* (plasmid 1) *qacA/B* (plasmid 1)
329	2	5:4	*blaZ* (plasmid 1) *fusB* (plasmid 1) *ant(4')-Ib* (plasmid 1) *qacA/B* (plasmid 1)
355	4	4:4	*fusB* (plasmid 1) *blaZ* (plasmid 1) *qacA/B* (plasmid 1) *ant(4')-Ib* (plasmid 1)
407	2	3:3	*mphC* (plasmid 1) *qacA/B* (plasmid 1) *msrA* (plasmid 1)
435	3	3:5	*ermC* (plasmid 1) *fusB* (plasmid 2) *AAC(6')-le—APH(2'')-la* (plasmid 2) *ANT(4')-Ib* (plasmid 2) *mupA* (plasmid 2)
436	3	8:1	*ermC* (plasmid 1)
465	1	7:0	No genes
475	5	4:2	*mupA* (plasmid 1) *ermC* (plasmid 2)
631	3	8:1	*ermA* (plasmid 1)
664	3	6:0	
673	3	11:2	*ermA* (plasmid 1) *tetK* (plasmid 2)
711	3	7:3	*Cat* (plasmid 1) *Ant(4')-Ib* (plasmid 2) *ermC* (plasmid 3)
712	3	7:3	Cat (plasmid 1) *Ant(4')-Ib* (plasmid 2) *ermC* (plasmid 3)
713	3	7:3	*Cat* (plasmid 1) *Ant(4')-Ib* (plasmid 2) *ermC* (plasmid 3)

S. epidermidis ID	Plasmids	Chromosome: plasmid-encoded antibiotic genes	Antibiotic resistance genes (and plasmid they are found on)
715	2	7:3	*Cat (*plasmid 1) *mupA* (plasmid 2) *Ant(4')-Ib* (plasmid 1)
716	4	7:3	*Cat* (plasmid 1) *Ant(4')-Ib* (plasmid 2) *ermC* (plasmid 3)

Table 4.8: Number of plasmids and the antibiotic resistance genes carried on plasmids found in *mecA*$^{+}$ *S. epidermidis* isolates recovered from East and West London

4.3.6 Phage prediction

Phage insertion was identified as a different level of completeness in 17 *S. epidermidis* isolates (Table 4.9). All the phage's insertions, which were considered to be complete, were phages that are associated with staphylococci. Interestingly, there was incomplete phage's insertion prediction with proteins (protease and integrases) that show homology to phages associated with *Bacillus; Planktothrix and, Streptococcus.* There were no antibiotic resistance genes identified within the phage insertion sequence. Of the 17 *mecA*$^{+}$ *S. epidermidis* isolates from public areas, 1 isolate (5.9%) had 4 phage insertion regions, 10 isolates (58.8%) had 3 phage insertion regions and 2 isolates (11.8%) that have 2 phage insertion regions and 4 isolates (17.7%) had only 1 phage insertion region. All intact *Staphylococcus phages* belong to the *Siphoviridae* family.

S. epidermidis ID	No phage	Most common phages	Completeness
321	3	*Staphylococcus phage StB12* *Staphylococcus phage phi5967PVL* *Staphylococcus phage CNPH82*	Intact Intact Incomplete
327	3	*Staphylococcus phage StB12* *Staphylococcus phage phi5967PVL* *Staphylococcus phage CNPH82*	Intact Intact Incomplete
329	3	*Staphylococcus phage StB12* *Staphylococcus phage phi5967PVL* *Staphylococcus phage CNPH82*	Intact Intact Incomplete
355	1	*Staphylococcus phage StB12*	Incomplete
407	3	*Staphylococcus phage* 187 *Staphylococcus phage CNPx* *Staphylococcus phage CNPx*	Incomplete Incomplete Intact
435	3	*Staphylococcus phage SPbeta-like* *Staphylococcus phage StauST398-2* *Streptococcus phage Dp-1*	Incomplete Incomplete Incomplete
436	2	*Staphylococcus phage StauST398-2* *Staphylococcus phage SPbeta-like*	Incomplete Incomplete
465	1	Staphylococcus phage PT1028	Incomplete
475	4	*Planktothrix phage PaV-LD* *Staphylococcus phage SPbeta-like* *Staphylococcus phage StauST398-2* *Streptococcus phage Dp-1*	Incomplete Incomplete Incomplete Incomplete
631	1	*Staphylococcus phage PT1028*	Incomplete
664	2	*Staphylococcus phage 187* *Staphylococcus phage CNPx*	Incomplete Questionable
673	1	*Staphylococcus phage PT1028*	Incomplete
711	3	*Staphylococcus phage PT1028* *Staphylococcus phage 187* *Bacillus phage vB_BhaS-171*	Incomplete Intact Incomplete
712	3	*Staphylococcus phage PT1028* *Staphylococcus phage 187* *Bacillus phage vB_BhaS-171*	Incomplete Intact Incomplete
713	3	*Staphylococcus phage 187* *Staphylococcus phage PT1028* *Bacillus phage vB_BhaS-171*	Intact Incomplete Incomplete
715	3	*Staphylococcus phage 187* *Staphylococcus phage 80* *Bacillus phage vB_BhaS-171*	Intact Incomplete Incomplete
716	3	*Staphylococcus phage PT1028* *Staphylococcus phage phi5967PVL* *Bacillus phage vB_BhaS-171*	Incomplete Intact Incomplete

Table 4.9: Phage prediction in *mecA*$^{+}$ *S. epidermidis* isolates from public areas from hospitals from East and West London

4.3.6 Phylogenetic analysis

Phylogenetic analysis was performed to determine the relatedness of environmental isolates in this study with those recovered from other sources, including clinical isolates. In this study, 17 *mecA*$^{+}$ *S. epidermidis* isolates recovered from East and West London were compared to WGS reference and draft isolates from the ENA database that was recovered from clinical samples (n=34); healthy humans (n=9), livestock (cows, pigs and sheep) and rodents (n=15), plant isolates recovered from rice (n=4), hospital environment from wards (n=7), livestock housing environment (n=2) and natural environment (n=2). *MecA* gene was present in 59 out of 90 isolates studied. From the core SNP phylogenetic tree, two distinctive clades were identified (Figure 4.2). Four isolates from East London hospital (321, 327, 329 and 355) belong in clade A whereas 1 East London isolate (407) was identified in clade B together with all (n= 12) West London isolates. Interestingly, all clinical isolates from ENA database except for VCU128 found in human airways were found to be in clade B. *S. epidermidis* isolates obtained from the ENA database recovered from human, animal and environmental were found on both clades where *S. epidermidis* isolates from the ENA database from plants (all rice seeds) were found only in clade A. Public setting isolates 355 were genetically related to ENA database isolates recovered from healthy humans (MRSE 52-2 and NIHLM057); isolate 407 was genetically related to isolates recovered from cow (Y24), pig (PR246B0) and animal housing (M01 and M025) and isolates whereas 435, 436, 465, 475, 631, 673, 711, 712, 713, 715 and 716 were genetically related to isolates from the ENA database recovered from clinical samples from blood and an endotracheal tube biofilm of a mechanically ventilated patient (ET-0240). 321 327 and 329 from East London Hospital were shown to be not closely related to any other isolate. The majority of the isolates that were phylogenetically related to isolates recovered from public setting in London carried the *mecA* gene. Interestingly,

there was a phylogenetic clustering of isolates, which were ST2 and ST5 but were broken up by isolates 435 (ST118) and isolates 631 and 673 (ST87) respectively.

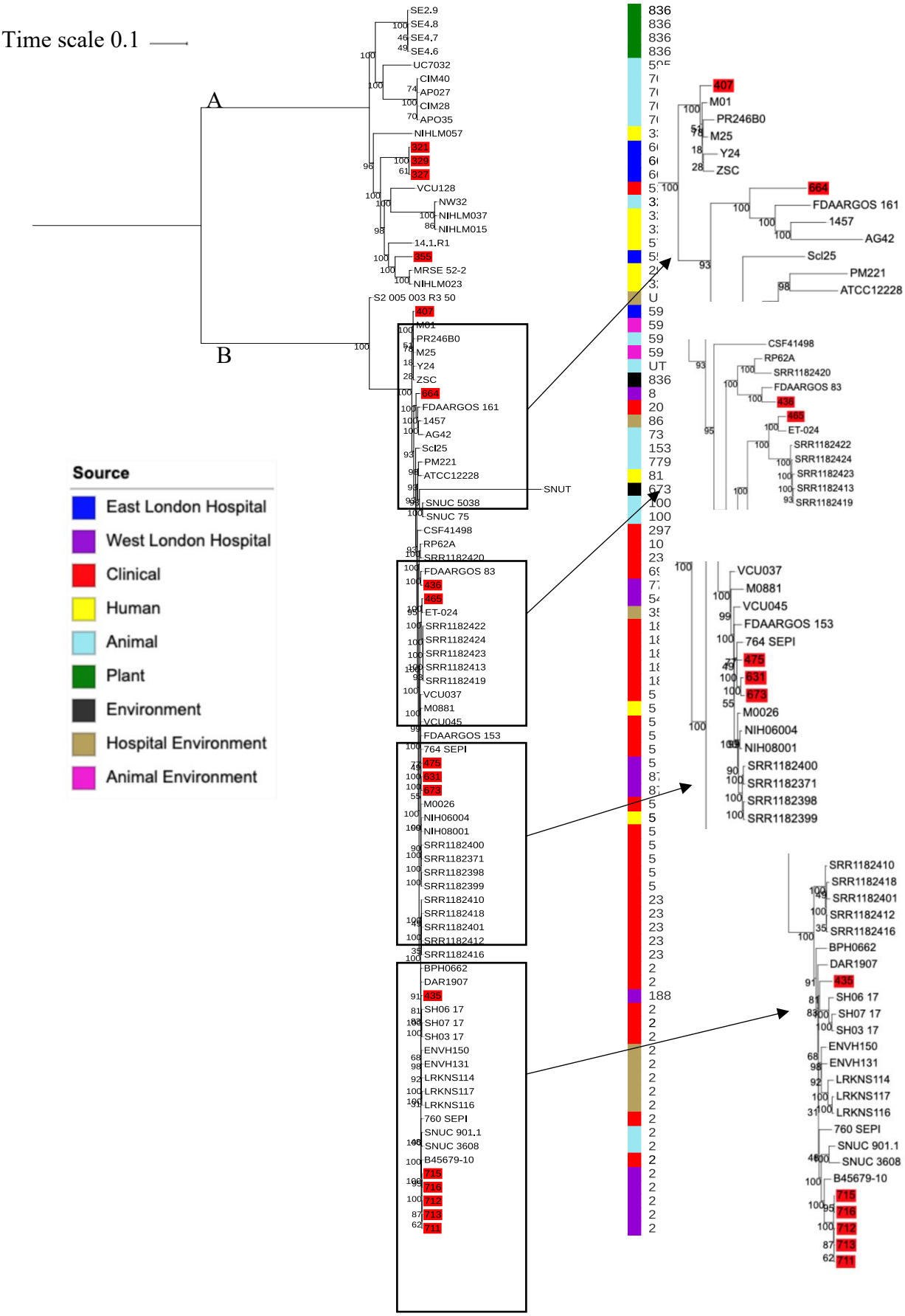


Figure 4.2: SNP core phylogenetic tree of 90 *S. epidermidis* isolates recovered from different sources. ST= MLST sequence type; UT=untypable. Red isolate background indicates

mecA^{+} *S. epidermidis* isolates from public settings in East and West London. Red Labels are *mecA*$^{+}$ isolates from public setting from East and West London.

PopPUNK analyses revealed there were 31 groups by their combined core and accessory genome. The accessory genome t -SNE analyses set at the perplexity of 20 showed that there were 5 distinct clusters with two groups showing a mixed group of isolates belonging to different combined clusters (Figure 4.3). No cluster had a single multilocus sequence type. In this study, *mecA*$^{+}$ *S. epidermidis* isolates recovered from East and West London were found in many different clusters. The accessory genome of *mecA*$^{+}$ *S. epidermidis* isolates was related to isolates from the ENA database recovered from clinical samples from blood, airways and cerebrospinal fluid, endotracheal tube biofilm of a mechanically ventilated patient and central venous catheter; healthy human skin, mucosa and airways; from livestock (cows, pigs and sheep), mouse, plants, and natural environment. Additionally, cluster 1 had isolates from the ENA database that were recovered from a hospital environment in medical wards and isolates recovered from clinical blood samples whereas cluster 4 are all isolates from clinical samples (blood) and a single isolate from healthy human skin (M008).

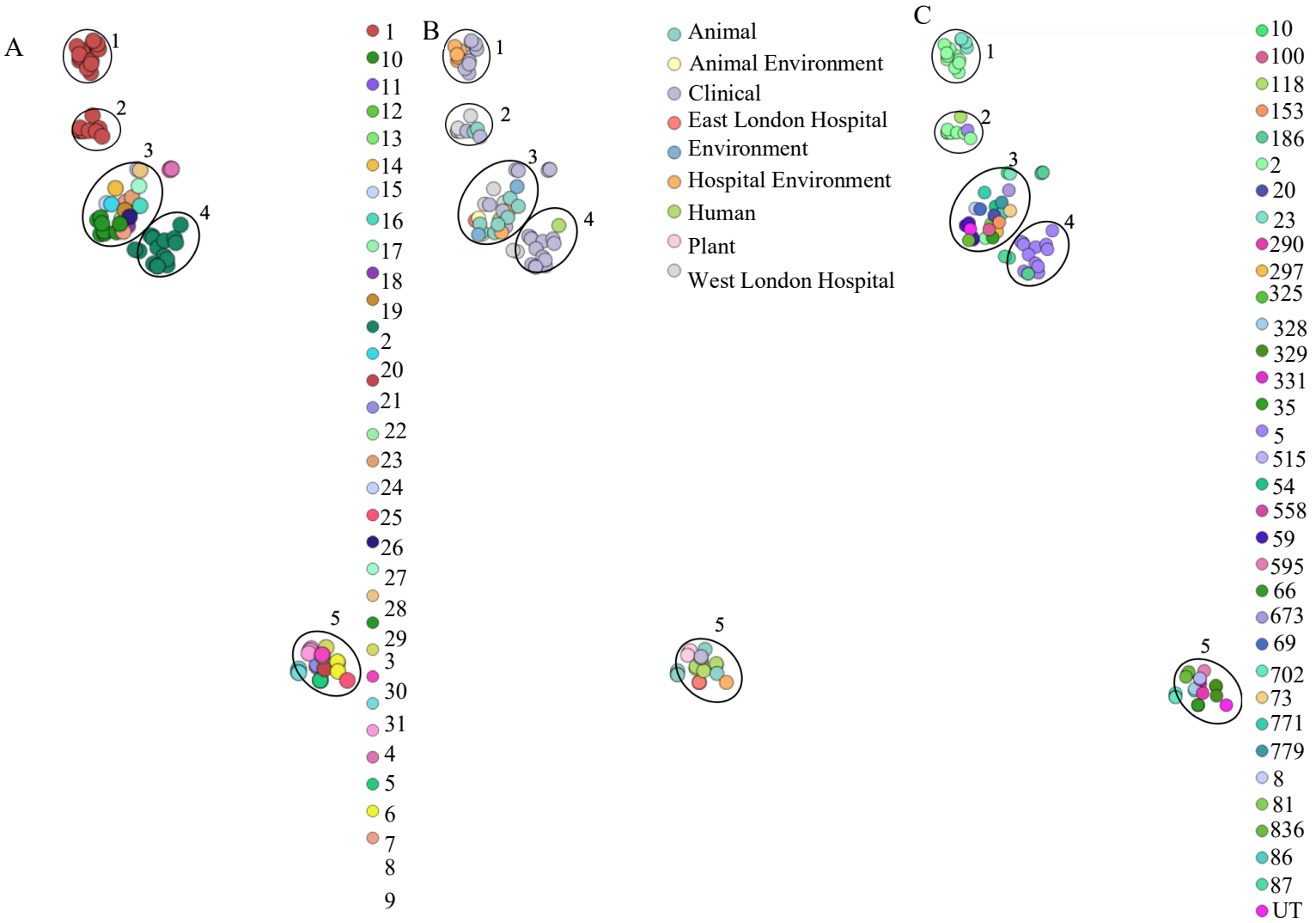


Figure 4.3: **t-SNE analyses of the distance of the accessory genome in 90 *S. epidermidis* isolates**. The analysis was performed using the PopPUNK pipeline. A maximum number of mixture components was set at 5 and for perplexity of t-SNE set at 20. (A) Combined cluster from PopPUNK analysis; (B) isolation source and (C) MLST.

4.3.7 Pangenomic analysis

A pangenome analysis was performed using the 90 *S. epidermidis* isolates used in the phylogenetic analysis including the 17 *mecA*$^{+}$ *S. epidermis* isolates from East and West London using the Roary pipeline (Page et al., 2015). In total, 8,590 genes were identified, of which 930 genes (10.8%) were core. 705 of the genes (8.2%) were soft-core genes (95 to 99% of isolates have these genes); 1,357 of the genes (15.8%) were shell genes (15 to 95% of isolates have these genes) and 5,598 genes (65.2%) were cloud genes (≤ 15%). From the pangenome analyse, a hierarchy clustering heatmap was constructed on the 8,590 genes that are present in the pangenome analyse to determine if isolates from the same source share the same gene profiles (Figure 4.4). With the *S. epidermis* isolate that were analysed there were no observed difference based on the hierarchy clustering heatmap except for isolates recovered from rice seed.

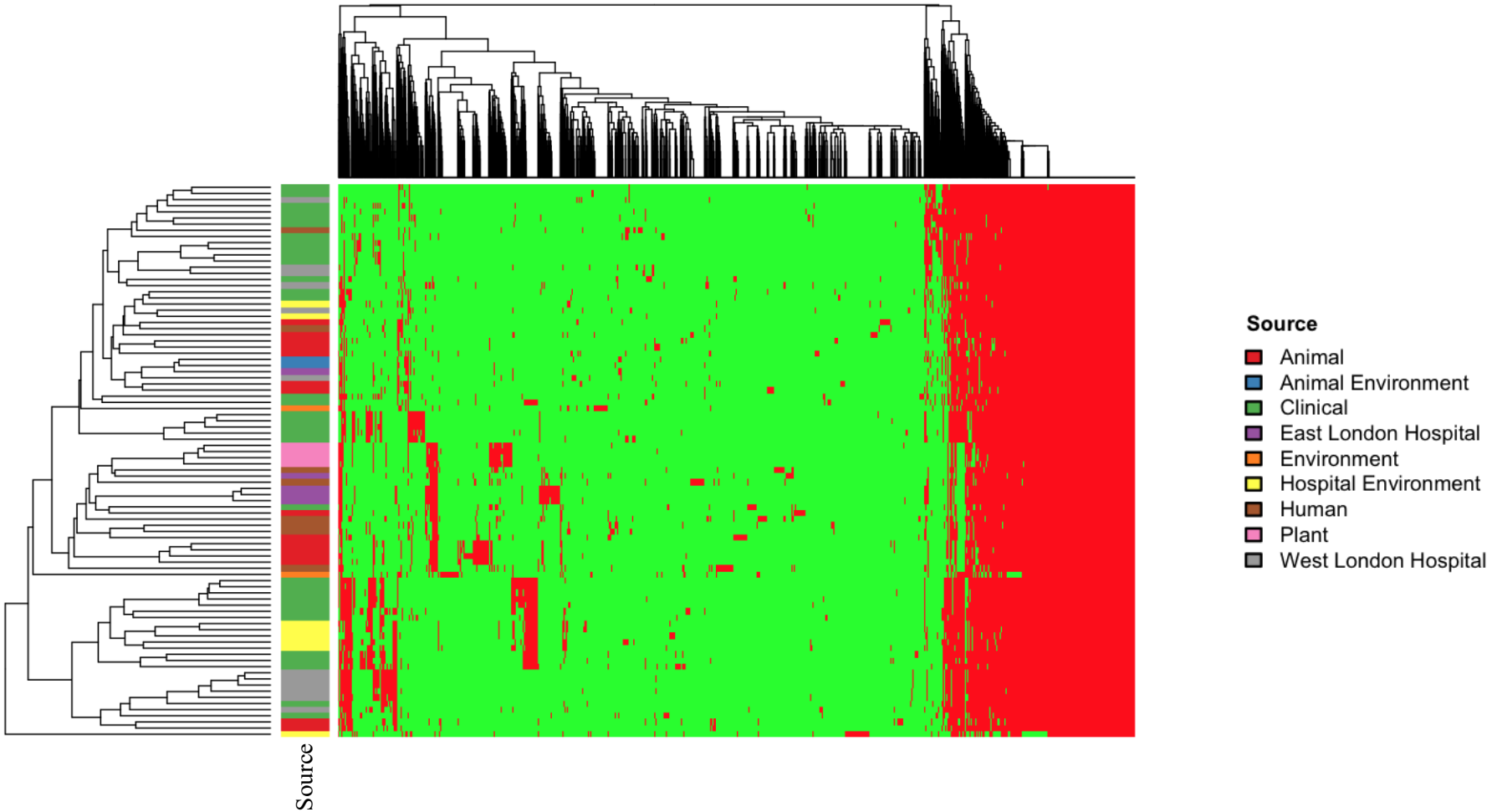


Figure 4.4: **Hierarchy clustering heatmap of presence and absence of all genes found in 90 *S. epidermidis* isolates from the pangenome analyse based on their isolation source.** Red tile represents the presence of gene; green tile represents the absence of the genes**.**

31 antibiotic resistance genes and 36 virulence genes were identified in the isolates in the pangenome analysis (Table 4.10 and Table 4.11). Interestingly, *norA* resistance genes and *geh* and *sspB* virulence genes were ubiquitous in these isolates. No antibiotic resistance genes or virulence genes were unique to *mecA*$^{+}$ *S. epidermidis* isolates in East and West London. Antibiotic resistance genes that were only found in a single isolate are *fosB3, fexA*, *APH(3')-IIIa*, *sat-4a*, *mepA*, *mepR*, *CTX-m-109* and *TEM-122*. The *vatB* gene was only present in isolates from West London. Virulence genes which were found in single isolates were *sdrC, sec and sell*. *FosB3*, *fexA* were found in isolate from the ENA database recovered from clinical ward; *APH(3')-IIIa*, *sat-4a* from a healthy human isolate and *mepA*, *mepR*, *CTX-m-109* and *TEM-122* from animals. *Sec* and *sell* virulence genes were identified in animal isolates and *sdrC* in a clinical isolate.

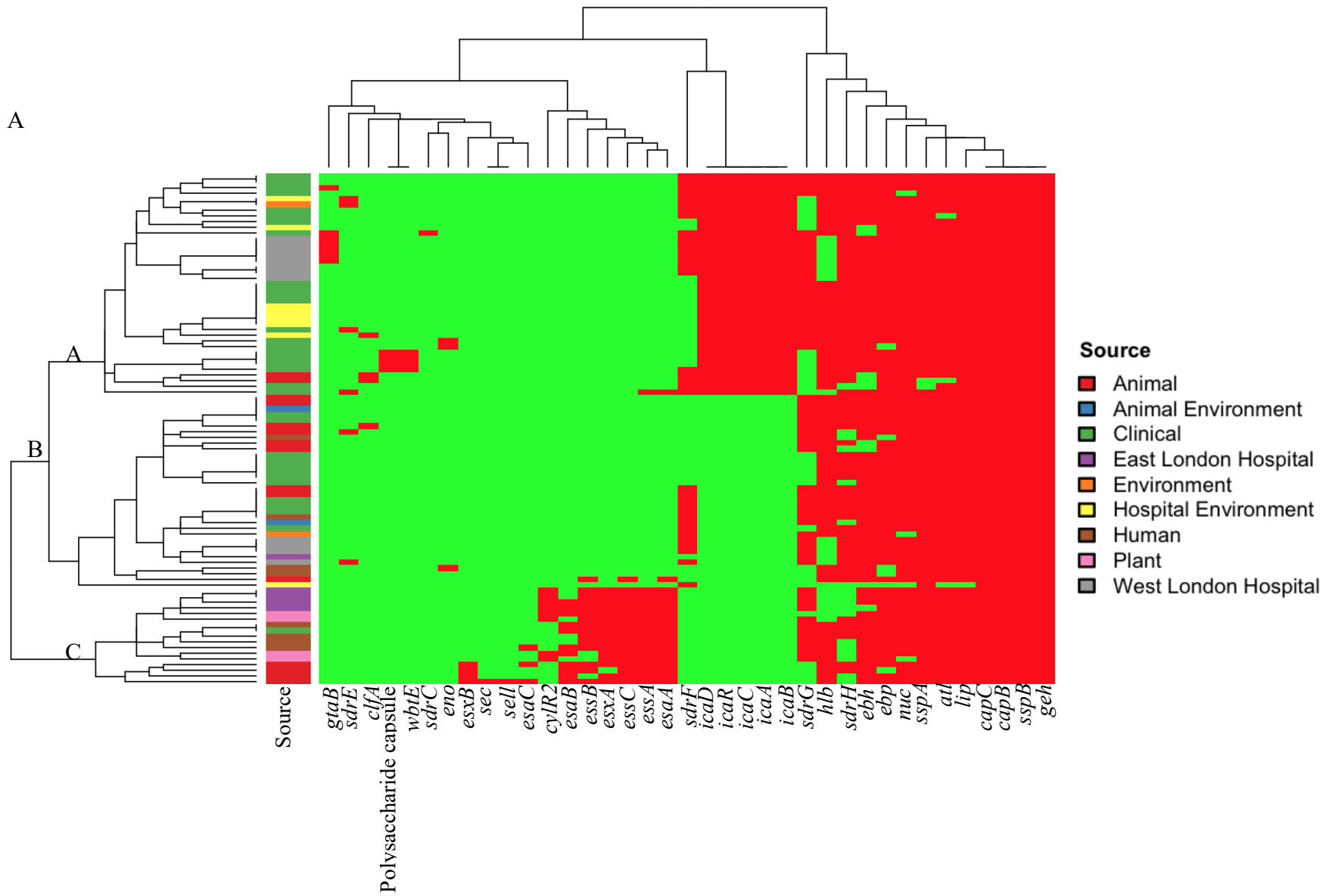
Antibiotic resistance Gene	List of antibiotic classes	Percentage (%)
mecA	Penam, monobactam, cephalosporin, carbapenem, cephamycin	65.6
norA	Fluoroquinolone, acridine dye	100.0
dfrC	Diaminopyrimidine	94.4
AAC(6')-Ie-APH(2'')-Ia	Aminoglycoside	44.4
qacA/B	Fluoroquinolone	48.9
blaZ	Beta-lactum	80.0
mupA	Muprcion	22.2
ermC	lincosamide, streptogramin, macrolide	25.6
mphC	macrolide phosphotransferase	16.7
msrA	streptogramin, macrolide	20.0
ANT(4')-Ib	aminoglycoside	25.6
tet(K)	Tetracycline	15.6
dfrG	Diaminopyrimidine	3.3
fusB	Fusidic acid	18.9
fosB3	Fosfomycin	1.1
fexA	Phenicol	1.1
cfrA	oxazolidinone antibiotic, streptogramin antibiotic, lincosamide antibiotic, phenicol antibiotic, macrolide antibiotic, pleuromutilin antibiotic	5.6
lnuA	Lincosamide	8.9
APH(3')-IIIa	Aminoglycoside	1.1
SAT-4A	Nucleoside antibiotic	1.1
mgrA	peptide antibiotic,penam,cephalosporin,acridine dye,fluoroquinolone, tetracycline	24.4
vgaA	streptogramin antibiotic, pleuromutilin antibiotic	2.2
mepA	tetracycline antibiotic, glycylcycline	1.1
mepR	tetracycline antibiotic, glycylcycline	1.1
CTX-m-109	cephalosporin	1.1
TEM-122	monobactam, penam, cephalosporin	1.1
vatB	streptogramin antibiotic	2.2
cat	phenicol	12.2
ermA	macrolide, streptogramin, lincosamide	3.3
isaB	pleuromutilin antibiotic, lincosamide antibiotic, streptogramin antibiotic	3.3

Table 4.10: Resistance genes in the *S. epidermidis* pangenome

Virulence gene	Function	Percentage (%)
atl	Autolysin	96.7
ebh	Cell wall-associated fibronectin-binding protein	90.0
ebp	Elastin binding protein	93.4
sdrF	Ser-Asp rich fibrinogen-binding proteins	38.9
sdrG	Ser-Asp rich fibrinogen-binding proteins	66.7
sdrH	Ser-Asp rich fibrinogen-binding proteins	80.00
sspB	Cysteine protease	100.0
geh	Lipase	100.0
lip	Lipase	98.9
sspA	Serine V8 protease	97.8
hlb	Beta hemolysin	76.7
icaA	Intercellular adhesion	43.3
icaB	Intercellular adhesion	43.3
icaC	Intercellular adhesion	43.3
icaD	Intercellular adhesion	43.3
icaR	Intercellular adhesion	43.3
nuc	Thermonuclease	95.6
esaA	Type VII secretion system	21.1
essA	Type VII secretion system	20.0
essB	Type VII secretion system	16.7
essC	Type VII secretion system	20.0
esxA	Type VII secretion system	17.8
cylR2	Cytolysin	8.9
sdrE	Ser-Asp rich fibrinogen-binding proteins	6.7
esaB	Type VII secretion system	12.2
esxB	Type VII secretion system	4.4
esaC	Type VII secretion system	3.3
eno	Streptococcal enolase	3.3
Capsule	Polysaccharide capsule	2.2
capB	Polyglutamic acid capsule	100
capC	Polyglutamic acid capsule	100
gtaB	Polysaccharide capsule	7.8
sdrC	Ser-Asp rich fibrinogen-binding proteins	1.1
wbtE	LPS	4.44
clfA	Clumping factor A	4.
sec	Enterotoxin C	1.1
sell	Enterotoxin-like L	1.1

Table 4.11: Virulence genes from 90 *S. epidermidis* isolates in the pangenome

A hierarchy clustering heatmap were constructed on the antibiotic resistance genes and virulence genes found in these 90 *S. epidermidis* isolates (Figure 4.5). There was no distinct clustering by source or genes for antibiotic resistance genes (Figure 4.5B). Interestingly, there was a distinct cluster which can be segregated by their virulence genes and isolation source (Figure 4.5A). Cluster A was mainly composed of isolates from ENA database recovered from clinical samples (blood and cerebrospinal fluid), natural environments, medical wards and a single livestock isolate (SNUC 3608) all of which had the *icaADBCR* operon, encoding the intercellular adhesion proteins for biofilm formation. This group contained *mecA*$^{+}$ *S. epidermidis* isolates 435, 436, 465, 711, 712, 713, 715 and 716 recovered from West London hospital. Cluster C was grouped by the *essA, essB, essC*, *esaA*, *esaB* and *esxA* genes, which encode for type VII secretion system. This cluster includes isolates from the ENA database recovered from rodents, rice seeds healthy human skin and airways as well as the *mecA*$^{+}$ *S. epidermidis* isolates 321, 327, 329 and 355 from East London hospital. Cluster B was grouped by the lack of the *icaADBCR* operon or the type VII secretion system genes. This cluster includes isolate from the ENA database recovered from clinical samples (blood and urine), livestock (cow, pig and sheep); animal housing; groundwater as well as the *mecA*$^{+}$ *S. epidermidis* isolates 407 from the East London hospital and isolate, 475, 631, 664 and 673 from the West London Hospital.


A
A
B
C
Source
gtaB
sdrE
clfA
Polysaccharide capsule
wbtE
sdrC
eno
esxB
sec
sell
esaC
cylR2
esaB
essB
esxA
essC
essA
esaA
sdrF
icaD
icaR
icaC
icaA
icaB
sdrG
hlb
sdrH
ebh
ebp
nuc
sspA
atl
lip
capC
capB
sspB
geh
Source
Animal
Animal Environment
Clinical
East London Hospital
Environment
Hospital Environment
Human
Plant
West London Hospital

B

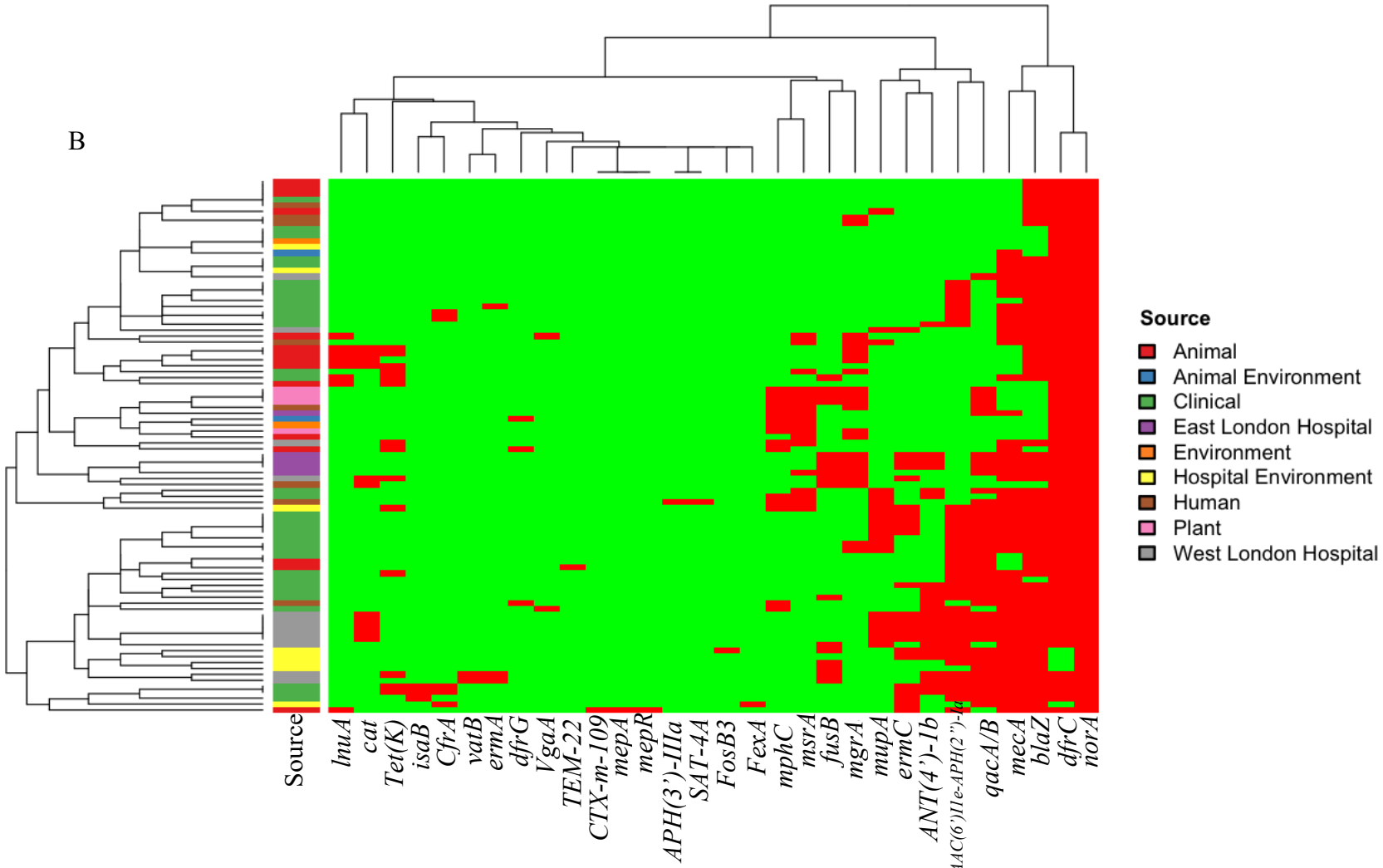


Figure 4.5: A hierarchy clustering heatmap analysis of 90 *S. epidermidis* isolates antibiotic resistance and virulence gene profiles based on their isolation source (A) Hierarchy clustered heatmap of antibiotic resistance genes found by source; (B) Hierarchy clustered heatmap of virulence gene by source. Red tile present; green tile absent.

COG family group was identified for core and accessory genome in *S. epidermidis* (Table 4.12). Interestingly the core genome showed to have a large portion of the genes which have general function prediction only (11.7%) and function unknown (11.1%). For the accessory genome, a large portion of the gene's function was for replication, recombination and repair (14.7%).

	Percentage (%) of genes in the pangenome	
Function	Core	Accessory
Translation, ribosomal structure and biogenesis	9.0	4.4
RNA processing and modification	0.0	0.0
Transcription	5.8	7.8
Replication, recombination and repair	4.1	14.7
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	1.0	1.8
Nuclear structure	0.0	0.0
Defence mechanisms	1.0	3.9
Signal transduction mechanisms	2.7	3.6
Cell wall/membrane/envelope biogenesis	4.0	5.1
Cell motility	0.1	0.2
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	1.1	1.2
Posttranslational modification, protein turnover, chaperones	3.5	2.3
Energy production and conversion	7.8	2.9
Carbohydrate transport and metabolism	6.7	6.1
Amino acid transport and metabolism	9.8	8.3
Nucleotide transport and metabolism	4.9	2.2
Coenzyme transport and metabolism	5.6	3.0
Lipid transport and metabolism	2.6	2.4
Inorganic ion transport and metabolism	5.9	6.3
Secondary metabolites biosynthesis, transport and catabolism	1.7	1.1
General function prediction only	11.7	12.9
Function unknow	10.9	9.8

Table 4.12: **COG family of the core and accessory genes in 90 *S. epidermidis* isolates used in pangenome analysis.**

The COG family that was identified for the unique genes found in only the 17 *mecA*$^{+}$ *S. epidermidis* isolates from public settings in East and West London was not found in the *S. epidermidis* isolates from the ENA database used in phylogenetic and pangenome analysis (Table 4.13). In total, there were 324 functional genes that uniquely found *mecA*$^{+}$

S. epidermidis isolates from public settings. Interestingly, 24.4% of these unique functional genes were identified to be for replication, recombination and repair

Gene Function	Percentage (%) of genes in the pangenome unique to S. *epidermidis* isolates from East and West London
Translation, ribosomal structure and biogenesis	3.8
Transcription	17.9
Replication, recombination and repair	24.4
Cell cycle control, cell division, chromosome	2.3
Defense mechanisms	7.5
Signal transduction mechanisms	1.2
Cell wall/membrane/envelope biogenesis	2.9
Intracellular trafficking, secretion, and vesicular transport	0.6
Posttranslational modification, protein turnover, chaperones	0.2
Energy production and conversion	2.8
Carbohydrate transport and metabolism	5.8
Amino acid transport and metabolism	1.6
Coenzyme transport and metabolism	2.2
Inorganic ion transport and metabolism	1.8
General function prediction only	11.3
Function unknown	13.7

Table 4.13: Unique gene COG family from 17 *S. epidermidis* isolates recovered from public settings in East and West London.

Comparative analysis of the accessory genes between *mecA*$^{+}$ isolates recovered from East and West London show that East London isolates had 579 genes that were not found in West London isolates whereas East London isolates had 678. For both East and West London isolates there was 1 gene that was found in all the isolates recovered from these areas. East London had a significantly higher portion of unique genes compared to West

London isolates for the COG function cell wall/membrane/envelope biogenesis compared with West London isolates (7.7% and 3.8% respectively (P= 0.0027)) and carbohydrate transport and metabolism (9.7% and 6.1% respectively (P=0.0158)). West London had a significantly higher proportion of unique genes compared to East London unique genes for the COG function replication, recombination and repair compared to East London isolates (18.3% and 13.4% respectively (P=0.0162)); signal transduction mechanisms (3.9% and 1.4% respectively (P=0.0059)) and amino acid transport and metabolism (8.9% and 5.8% respectively (P=0.0403)) (Table 4.14).

	Percentage (%) of genes in the pangenome	
Function	East London	West London
Translation, ribosomal structure and biogenesis	2.9	2.7
RNA processing and modification	0.0	0.0
Transcription	9.9	9.4
Replication, recombination and repair	13.4	18.3*
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	1.0	1.5
Nuclear structure	0.0	0.0
Defence mechanisms	4.5	6.6
Signal transduction mechanisms	1.4	3.9*
Cell wall/membrane/envelope biogenesis	7.7*	3.8
Cell motility	0.1	0.2
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	1.3	1.9
Posttranslational modification, protein turnover, chaperones	0.3	1.3
Energy production and conversion	2.1	1.0
Carbohydrate transport and metabolism	9.7*	6.1
Amino acid transport and metabolism	5.8	8.9*
Nucleotide transport and metabolism	1.9	1.8
Coenzyme transport and metabolism	3.1	2.1
Lipid transport and metabolism	2.6	1.3
Inorganic ion transport and metabolism	7.8	8.2
Secondary metabolites biosynthesis, transport and catabolism	2.8	1.4
General function prediction only	13.0	11.1
Function unknow	8.6	8.5

Table 4.14: **COG function of genes that were uniquely found in from *mecA*$^{+}$ *S. epidermidis* isolates from East London and *mecA*$^{+}$ isolates from West London.***= Area where there was significantly higher (p=<0.05) percentage of genes for that COG function.

4.4 Discussion

4.4.1 MLST of *mecA*$^+$ *S. epidermidis* isolates recovered from East and West London

MLST is still widely used in pathogens to study population genetics and evolution (Thomas and Robinson, 2014). MLST data from *mecA*$^+$ S. *epidermidis* isolates from public settings in East and West London showed a wide range of genetic variability. Consistent with previous reports studying ST2 was the most common sequence type identified in this study (Deplano et al., 2016; Dong et al., 2018; Widerström et al., 2012). In this study isolates that harboured ST2 sequence types were isolated from public areas in hospitals. In addition, in this study, two new sequence types designated as ST771 and ST779 were identified in isolates recovered from a high frequency touched surfaces from a hospital in West London. These results show that are genetic lineages of *S. epidermidis* currently not known about which have the potential to spread antibiotic resistance genes within hospitals areas.

4.4.2 Virulence genes detected in *mecA*$^+$ *S. epidermidis* isolates from public areas in East and West London

S. epidermidis has previously been shown to have a large array of virulence genes. One of the key virulent factors that have been identified is its ability to produce biofilm which helps *S. epidermidis* adhere to abiotic surfaces (Büttner et al., 2015; Cho et al., 2002; Trampuz and Zimmerli, 2005). This virulent factor is encoded by the *icaADBC* and the *icaR* regulator (Cho et al., 2002). 47.1% of *mecA*$^+$ isolates from hand-touched surfaces in this study, possessed *icaADBC* operon. Similar data were reported in a hospital from Poland which found that 46.9% of the *S. epidermidis* isolates recovered from air and surfaces from surgical theatres and general surgical wards had the *icaADBC* operon (Wojtyczka et al., 2014). Though it was not determined if public setting *S. epidermidis*

isolates have the biofilm phenotype, it was still considered that isolates which have the biofilm-producing gene to be a public health risk. Whereas the isolates which lack these genes would be considered to be less virulent as biofilm production has been linked to *S. epidermidis* ability to adhere to host cells, invade host immune system as well as being more tolerant to several classes of antibiotics (Ghasemian et al., 2012; Kristian et al., 2008; Otto, 2008)

From this study, the *mecA*$^{+}$ *S. epidermidis* isolates from public settings were found to carry genes associated with the type VII secretion system. These virulent factors have been characterised in *S. aureus* to secrete a nuclease toxin that targets competing bacteria as well as protects them against antimicrobial host fatty acids, but this has not been reported in *S. epidermidis* (Cao et al., 2016; Tchoupa et al., 2019). The type VII secretion system is encoded by cluster of 12 genes of which four of the genes encode for the membrane-associated proteins (*esaA*, *essA*, *essB*, and *essC*); three genes encode for the cytosolic proteins (*esaB*, *esaE*, and *esaG*), and five genes encode for the secreted virulence factors (*esxA*, *esxC*, *esxB*, *esxD*, and *esaD*) (Tchoupa et al., 2019) Interestingly, 6 out of the 12 type VII secretion system genes were identified in the isolates recovered from East London hospital. The 6 genes that were not present in these isolates were two genes that encode for the cytosolic proteins (*esaE* and *essG*) and four genes that encode for the secreted virulence factors (*esxC*, *esxB*, *esxD*, and e*saD*). Additionally, isolate 321, 327 and 407 from East London hospital also lacked the *esaB* gene. Interestingly, all these East London hospital isolates all carried the genes which encode for membrane-associated proteins. Isolates that lack one of *esxA* or *esxB* genes were shown to have a reduced virulence (Burts et al., 2008, 2005). This would suggest that these isolates will not be able to produce a virulence factor, but it is worrying to see these genes in the first place as it

is possible that all genes that are required to cause virulence could be horizontally transferred to *S. epidermidis.*

Two genes which were identified in 100% of these *mecA*$^{+}$ *S. epidermidis* isolates recovered from public setting in East and West London was *capB* and *capC*. These genes are normally found on the locus with two other genes, *capA* and *capD* (Otto, 2012). Currently, there are no studies that describe the lack of these genes from the *cap* locus would effects the production of the polyglutamic capsule. Interestingly, *capB* and *capC* were ubiquitous in isolates included in the pangenome analyses; however, *capA* or *capD* were not identified. This may be due to lack of references for *capA* and *capD* sequences to get a sufficient alignment.

Polysaccharide capsule gene *gtaB* usually associated with bacillus species virulence was identified in 29.4% of the *mecA*$^{+}$ S. *epidermidis* isolates recovered from public settings in East and West London. This gene has previously been identified in *S. epidermidis* in sugar metabolism but has not been reported as a virulence factor (Gründling and Schneewind, 2007). Therefore, this would suggest that these *S. epidermidis mecA*$^{+}$ isolates from public settings have a *gtaB* gene is homolog to that of the *Bacillus* species *gtaB* gene but has evolved to have a different function. Another noticeable virulence gene in these isolates was *cylR2* which was predicted to be encoded in 23.5% of the *S. epidermidis* isolates. *CylR2* is a virulence gene regulator for the cytolysin operon found in *Enterococcus* spp. but not in staphylococci (Rumpel et al., 2004). Additionally, this gene is regulator as part of an 8 gene operon that encodes the cytolysin protein (Shankar et al., 2004). Cytolysin toxin has not been previously reported in *S. epidermidis* but has been reported in *S. aureus,* therefore, could be horizontally transferred across (Queck et al., 2009). As there is no virulence phenotype data in this study, it is not possible to determine if these virulence genes are functional.

4.4.3 Horizontal transfer of antibiotic resistance and virulence genes

HGT is an essential mechanism for bacterial survival (Polz et al., 2013). It has been shown that bacteria can acquire antibiotic resistance genes and virulence genes via transformation, conjugation and transduction. In this study, 17 *mecA*$^{+}$ *S. epidermidis* isolates recovered from public areas in East and West London Hospitals were analysed for potential HGT, in particular, to evaluate plasmid and prophage insertion into their genomes.

HGT was predicted to range from 8.6% to 10.8% of the total genome of the isolates from East and West London. In comparison, this was similar to that of the reference *S. epidermidis* genome ATCC 12288 (8.6%) from healthy human skin and mucosa used in this study though it was less than *S. aureus* isolates (15-20%) (Lindsay, 2010). For *S. epidermidis* there have been no previous genomic studies on the proportions of their genome that may have been horizontally transferred.

From the 17 *mecA*$^{+}$ *S. epidermidis* isolates it was predicted that a large portion of the horizontally transferred genes was predicted to be transferred from *Bacillus*, *Macrococcus* and *Salinicoccus* genus. These three genera are quite closely related to *Staphylococcus* as they belong to the same taxonomy order (bacillales) and *Macrococcus* and *Salinicoccus* are the same taxonomy family (*Staphylococcaceae*) (Becker et al., 2014; Fritze, 2004; Kumar et al., 2015; Kwok, 2003).

Interestingly, *the mecA* gene in all isolates was predicted to be horizontally transferred from *Macrococcus canis,* a bacteria which causes infection in dogs (Mašlaňová et al., 2018). Previous studies suggest that *mecA* gene and surrounding chromosomal region originated from *S. fleurettii* genome, although other studies found the *mecA* gene to be

encoded on a plasmid in *Macrococcus caseolyticus,* an ancestor of staphylococci (Baba et al., 2009; Tsubakishita et al., 2010). The HGTector tool explains that the predicted donor species should be considered as a "donor link" to describe the direction of gene transfer and the relationship between organisms (Zhu et al., 2014). Therefore, it is possible that the *mecA* gene was donated from *M. caseolyticus* as they have phylogenetically been shown to be close relatives to *M. canis* (Mašlaňová et al., 2018).

Antibiotic resistance genes and virulence genes, which were found in different *mecA*$^{+}$ *S. epidermidis* isolates recovered from public setting in East and West London to have different donors. For example, there were two different predicted donor species for antibiotic resistance genes *ermC* in isolate *AAC(6')-Ie-APH(2'')-Ia,* and *mupA* and virulence gene *icaR and hlb*. This would suggest that these genes were acquired from different ancestors. From the virulence genes that were predicted, not all the genes were found within the *icaADBC* operon. In addition, the genes that encode type VII secretion system machinery were predicted as being horizontal transfer. Only *icaA* of the *icaADBC* operon and *essC* and *esxA* of type VII secretion system were predicted to be horizontal transferred. *IcaA* was predicted to be donated from a biofilm-producing *Macrococcus caseolyticus;* a close relative of staphylococci and been previously shown to have DNA fragments overlapped between the two genera (Mašlaňová et al., 2018). For *essC* and *esxA* gene, the predicted donors were *Listeria booriae* and *Bacillus* spp. *UNC41MFS5* respectfully. There were no previous reports of the two species to have type VII secretion genes before, but they were closely related to known pathogenic species like *Bacillus subtilus* and *Listeria monocytogenes* (Simeone et al., 2009). The method that was used for HGT detection was reported to have an 81.6% precision in predicting genomes. Errors in HGT detection can come from database errors and incompleteness, ambiguity in the

genomic makeup and phyletic pattern and the difficulties of phylogenetic reconstruction (Zhu et al., 2014).

Plasmids were identified in all isolates which ranged from having 1 to 4 plasmids of which 0 to 6 antibiotic resistance genes were carried on these isolates' plasmids. All the antibiotic resistance genes found to be carried on the plasmid in *S. epidermidis* have been identified in other *S. epidermidis* isolated before as reported on the comprehensive antibiotic resistance database except for *fusB* (Jia et al., 2017). The *fusB* gene has previously been reported to be carried on the *S. aureus* plasmid pUB101(O'Brien et al., 2002). This might suggest that these *S. epidermidis* isolates may have acquired the *fusB* gene by conjugating with *S. aureus* (LaBreck et al., 2018). Antibiotic resistance genes were both identified to be found a plasmid and predicted as being present within a genomic island in the same isolates. This is due to the software detecting plasmid DNA as foreign compared to the rest of the genome, based on the dinucleotide biased in 8 genes or more, identification of a mobility gene and/or codon usage ((Bertelli et al., 2017; Bertelli and Brinkman, 2018; Waack et al., 2006). Genes that were predicted to be part of a genomic island have always encoded on a plasmid was *blaZ*, *vatB*, *vgaB*, *qacA/B*, *AAC(6')-Ie-APH(2") -Ia*, *mupA* and *mecA*. From these isolates, the *mecA* should be able to identify within a genomic island, but only 4 of the 17 isolates were correctly identified. However, the *mecA* gene was near predicted genomic islands. This would suggest that the software identifies *mecA* and some genes nearby as being native to *S. epidermidis*. Interestingly, little is reported about resistance island in *S. epidermidis* except for SCC*mec* element (Barbier et al., 2010; Seng et al., 2017b; Xu et al., 2015). Further studies into genomic islands would help better understand the epidemiology of *S. epidermidis* and from where they might have acquired antibiotic resistance and virulence genes.

In the *S. epidermidis* isolates from public settings, there was no detectable antibiotic resistance gene or virulent genes within prophage insertion regions. This shows that the transfer of antibiotic resistance genes in phages in *S. epidermidis* is rare (Deghorain and Van Melderen, 2012). However, phage sequences that show homology to phages that are not associated with staphylococci. These incomplete phages insertion sequence were shown to have homology to phages associated with *Bacillus*, *Planktothrix* and *Streptococcus*. Different phages have shown to have a range of host they can affect from a single strain, species, genus or even bacteria in different genera (Ross et al., 2016). There are no known phages that can affect *Staphylococcus* and other bacterial genera. Therefore, it is most likely that these sequences are from uncharacterised *Staphylococcus* phages. Of the intact phages inserted into public setting *S. epidermidis* isolates the most interesting were *Staphylococcus* phage STB12 which has only been previously reported before in *S. hominis* and *S. capitis* isolates, *Staphylococcus* phage 187 which been identified to infect *S. aureus* and *Staphylococcus phage phi5967PVL* which also infects *S. aureus* and has previous shown to transfer the *lukS-PV* and *lukF-PV* genes which encodes for Panton-Valentine leukocidin (PVL) toxin in *S. aureus.* These genes were not found in the *mecA*$^{+}$ *S. epidermidis* isolates from public settings. (Daniel et al., 2007; Deghorain et al., 2012; Gutiérrez et al., 2012; M. Zhang et al., 2011). These findings show that *Staphylococcus* phages can infect and integrated into genomes across the genus. Therefore, it is possible for a reciprocal exchange of phages related to *S. aureus* and CoNS. This could indicate that these *S. epidermidis* isolates from public settings could acquire staphylokinases, superantigens, PVL virulent genes that are horizontally transferred between *S. aureus* strains via phages (Goerke et al., 2009). Acquiring these genes will likely make *S. epidermidis* more virulent and a more significant risk to public health.

4.4.4 Phylogenetic comparison of *S. epidermidis* isolates recovered from public areas in East and West London with reference *S. epidermidis* isolates from ENA database.

Previous phylogenetic studies for the core genome *S. epidermidis* have shown that the majority of isolates recovered from healthy humans, clinical isolates, sheep and some rodents can be found within one clade whereas the majority of isolates recovered from rodents; rice seeds and some isolates recovered from healthy humans can be found within a separate clade (Chaudhry and Patil, 2016; Conlan et al., 2012). In this study, there was a similar observation except for one isolate recovered from a clinical sample from airways (VCU128) was found in a separate clade to the other *S. epidermidis* isolates recovered from clinical samples (blood, cerebral spinal fluid and urine).

Isolates from the ENA database recovered from cows, pigs, sheep and their housing were found in a clade where the majority of the isolates were associated with humans. The *mecA*$^{+}$ *S. epidermidis* isolates from public areas from East and West London hospitals occupy in both clades showing they are genetically diverse. The majority of these isolates recovered from public settings in West London hospitals (711, 712, 713, 715 and 716) were shown to be genetically related to clinical isolates recovered from blood, by their core genome as well as their MLST (ST2) which has been reported to be one of the most common sequence types found in hospital-acquired infections (Deplano et al., 2016). This shows that strains that cause infection on medical wards can be found in public areas in hospitals. *MecA*$^{+}$ isolate 435, 475, 631 from West London hospital to be genetically related to clinical isolates from blood; isolates 436 from west London was genetically related to isolate from urine (FDAARGOS-83) and isolate 465 was genetically related to isolate recovered from endotracheal tube biofilm of a mechanically ventilated patient (ET-024). Additionally, some isolates from the ENA database were recovered from livestock, and their housing was genetically related to an isolate which was recovered

from public areas in hospitals (407). These isolates belong to the same MLST (ST59), which has previously been reported to be both livestock and human-associated sequence type (Argudín et al., 2015). The *mecA*$^{+}$ public setting isolates were genetically related to ENA database isolates from cows which have bovine mastitis (Y24). These findings indicate that isolates found in public areas in hospitals could also cause blood infections (septicaemia), urinary tract infection in humans and cause bovine mastitis in cows due to them being in the same genetic lineage of isolates that have caused these infections. Other studies have shown *mecA*$^{+}$ *S. epidermidis* to be a common cause in bovine mastitis and been recovered from cows milk (Fernandes Dos Santos et al., 2016; Feßler et al., 2010). Additionally, reports have shown that pigs are also a reservoir of *mecA*$^{+}$ *S. epidermidis* and there is an indication of strains exchanging between human and pigs by their virulence and antibiotic resistance gene profile (Argudín et al., 2015; Tulinski et al., 2012). These reports, combined with data from this study, suggest that *S. epidermidis* is zoonoses (spread between animal and humans) and that livestock *mecA*$^{+}$ *S. epidermidis* isolates are from the same genetic lineages of the isolates shown to cause infections in humans. It is possible that some known isolates that can cause infections in humans originate from animals and have been transferred to humans and their associated environment either via direct contact with farmers or via food.

The t-SNE accessory genome analyses show that 2 out of the 5 accessory clusters isolates evolutions were different from the predicted combined core and accessory genome. These two clusters also have a collection of isolates recovered from different sources where all clusters were shown to have many different MLST. This would suggest that isolates that have similar core genomes have different accessory genome due to horizontal transfer of genes from other organisms that might be present in different isolation sites. Additionally, it was observed that all the clusters in *S. epidermidis* t-SNE analyse had a mixture of

isolates that were recovered from many different sources. Though it was observed that a cluster (cluster 1) had isolates from the ENA database that were recovered from medical wards and from clinical samples and another cluster (cluster 3) to have isolates that were only recovered from clinical samples or healthy human skin. This would suggest that cluster 1 isolates are only present within hospitals and have not spread outside clinical settings and cluster 3 isolates only associated with humans.

MecA^{+} isolates that were recovered from public areas in hospitals were shown to be similar in the accessory genome to isolates from the ENA database recovered from clinical samples (blood, urine and cerebrospinal fluid) from livestock (cow, pigs and sheep), mouse, plants, and natural environment. This suggests that *MecA*$^{+}$ *S. epidermidis* isolates from public setting likely originated from bacteria species associated with a particular niche that has been transmitted to new niche via humans or via food in *S. epidermidis*. Public areas in hospitals could be a place where isolates from different areas can mix and horizontally transfer genes to each other from bacteria usually associated with a particular environment. Additionally, virulent factors are typically associated as part of the accessory genome in *Staphylococcus* (Bosi et al., 2016). As findings from this study show that isolates from public settings, livestock and plants were genetically related in their accessory genome to that of *S. epidermidis* isolates that have been shown to cause septicaemia, urinary tract infection and meningitis it would suggest isolates recovered from these environments may also have the potential to cause this infection.

4.4.5 Pangenome and comparison of antibiotic resistance and virulence genes of isolates recovered from public settings with isolates obtained from the ENA database

Pangenome analysis showed that *S. epidermidis* has a minimal core genome (10.8% of the total number of genes) in comparison to *S. aureus* (19.3% of the total number of genes)

(Bosi et al., 2016). Interestingly, there was a large number of genes (65.2%) that were only present in ≤15% of the isolates. This suggests that *S. epidermidis* isolates have an open pangenome with many variable genes. There was no predicted clustering by presence/absence of genes except for the rice samples. This would suggest that isolates from East and West London or other sources in this study do not have a gene or genes found in the accessory genome required for survival in a particular niche.

In pangenome analysis, *norA* resistance gene and *geh* and *sspB* virulence genes were found to be part of the core genome. *NorA* is an efflux transporter that can actively pump out quinolone antibiotics, whereas *geh* encodes for lipase which breaks down fatty acids, and *sspB* encodes for cysteine protease, which is able to breakdown elastin, fibronectin and kininogen (Cadieux et al., 2014; Massimi et al., 2002; Yu et al., 2002). It is possible that these genes are essential for *S. epidermidis* survival and may well employ roles other than antibiotic resistance and virulence. Interestingly, *S. epidermidis* can be put into three distinct groups by their virulence gene profile. One group had the *icaADBCR* operon for biofilm production but lacked the type VII secretion genes; the second group had the type VII secretion genes but lacked the *icaADBCR* operon and the final group lacked genes for both of these virulent factors. Further laboratory studies would be required to understand why the *S. epidermidis* isolate cannot have both the genes that encode biofilm production and the type VII secretion system.

From this study, there were a few antibiotic resistance genes and virulence genes that were identified that have not been reported before in *S. epidermidis*. *IsaB* antibiotic resistance gene has previously been reported in *S. sciuri* where *CTX-m-109* and *TEM-122* have only been reported to be present in Gram-negative bacteria (Kaye et al., 2004; Kehrenberg et al., 2004; Zhang et al., 2011). Both CTX-*m-109* and *TEM-122* encode for alternative beta-lactamase (Kaye et al., 2004; Zhang et al., 2011). These last two genes

were found in isolates recovered from livestock, suggesting that these genes may be unique to this environment. The *sdrC* virulence gene which encodes for the Ser-Asp rich fibrinogen-binding proteins has not been reported before in *S. epidermidis*, however, the known *S. aureus* virulent gene *sell* which encodes for Enterotoxin-like L and *sec* which encodes for the Enterotoxin C have been reported in few cases in *S. epidermidis* (Barbu et al., 2010; Madhusoodanan et al., 2011). The two enterotoxin genes have been identified to be part of the pathogenicity island SePI were previously identified in clinical *S. epidermidis* isolate FRI909 (Argemi et al., 2018). Though these virulence factors were not found in isolates recovered in this study, it is worrying to see virulent genes associated with *S. aureus* which could transform *S. epidermidis* to possess more aggressive virulence determinants contributing to its spread to other environments.

COG analyses of the core and accessory genome showed there was a large portion of the accessory genome genes responsible for replication, recombination and repair (14.8%). This was also noted from the unique genes that were only found in the *mecA*$^{+}$ *S. epidermidis* isolates recovered from public areas from East and West London (24.4%). The large percentage of *S. epidermidis* isolates genomes from London hospitals that were detected as HGT genes suggest that these unique genes are essential for acquiring mobile genetic elements as they may contribute to the survival of these organisms on high-frequency touched surfaces which is not their usual and preferred niche to live on. Additionally, in these environments, these isolates might experience reactive oxygen species (ROS) that can be found in disinfectants that can cause DNA damage to the bacteria cell by damaging protein and DNA (Sheng et al., 2015). Genes that are members of this COG group have previously been shown to help repair damage caused by ROS and protect the bacteria from disinfectants (Gaupp et al., 2012). There was a significant difference between a few of the COG functions found between S. *epidermidis* isolates

from East London hospital and West London hospital. As these isolates were recovered from similar areas in hospitals, it was expected that these differences are mainly due to horizontal transference of genes from bacteria found in that area than adaption required to survive that particular niche.

4. 5 Conclusion

In this chapter, public setting *mecA+ S. epidermidis* isolates from East and West London was genetically compared to *S. epidermidis* isolates recovered from clinical samples (blood, cerebrospinal fluid, urine); healthy humans skin, nares and nasopharynx; livestock (cows, pigs and sheep); rodents (rats), plants, hospital environment from wards and catheters, animal housing and natural environment from the ENA database. From phylogenetic analysis, the genetic relatedness of *S. epidermidis* isolates from public settings was determined to that of other isolates from the ENA database. Additionally, the antibiotic resistance genes, virulent genes, mobile elements and pangenome were compared. From these analyses, there were novel findings. These findings were:

1. *S. epidermidis* carried part of Type VII secretion system operon found *S. aureus*, which has not been previously characterised in *S. epidermidis*. Additionally, Type VII secretion genes were only found in isolates lacking the *icaADBC* operon which encodes for biofilm production.

1. *MecA*$^{+}$ *S. epidermidis* isolates from public areas were shown to carry novel genetic elements including phages that only been reported in *S. aureus* or other CoNS isolates. This shows *S. epidermidis* has the potential of acquiring virulence determents from *S. aureus,* which will make *S. epidermidis* more virulent.
2. Many of the *mecA*$^{+}$ *S. epidermidis* isolates from high-frequency touched surfaces from East and West London belong to same genetic lineages of isolates from the ENA database that can cause blood and urine infection which suggests these

isolates are a public threat. Additionally, the *mecA*$^{+}$ *S. epidermidis* isolates from public areas in hospitals accessory genome to be related to isolates from the ENA database that were recovered from clinical infection associated with septicaemia, urinary tract infection and meningitis; livestock (cows, pig and sheep) healthy humans and natural environment. This suggests that these areas of the hospital are a mixing area of different genetic lineages of *S. epidermidis* in which they could horizontally transfer genes that may be associated to a particular niche to another *S. epidermidis* isolate.

3. *MecA*$^{+}$ isolates from public setting had more unique genes compared to isolates from ENA database that encode for replication, recombination and repair. This may be due to ROS that can be found in disinfectants that can cause DNA damage to the bacteria cell by damaging protein and DNA.

Chapter 5: Comparative genomics of *mecA* positive *S. haemolyticus* isolates recovered from public settings using the One Health approach; determining horizontal gene transfer in these isolates

5.1. Introduction

Staphylococcus haemolyticus is an emerging opportunistic pathogen, primarily isolated from neonatal patients and care units (Pereira et al., 2014). Like *S. epidermidis*, *S. haemolyticus* can produce biofilm (Fredheim et al., 2009). Additionally, *S. haemolyticus* can produce a capsule similar to *S. aureus*, which protects the bacteria from phagocytosis (Flahaut et al., 2008). Previous studies have demonstrated this bacterium recovered from both clinical and environmental sources can harbour multidrug resistance genes (Barros et al., 2012; Seng et al., 2017b; Xu et al., 2015). However, little is known about the genetic diversity and molecular epidemiology of *S. haemolyticus* recovered from general public settings.

5.2 Method

In this chapter, comparative genomic analyses were performed on the *mecA*$^+$ *S. haemolyticus* isolates from public settings from East and West London. This included identification of horizontal gene transfer (HGT) within the genome and the potential donor organisms of horizontally transferred antibiotic resistance and virulence genes. Additionally, it was investigated if virulence or antibiotic resistance genes were transferred via plasmids, phages or by other mobile genetic elements. A phylogenetic tree was constructed to determine the genetic relationship of *mecA*$^+$ isolates from this study with *S. haemolyticus* isolates using reference and draft WGS isolates from the European Nucleotide Archive (ENA) database that were previously recovered from different sources, including isolates recovered from clinical samples (eye, blood, sputum, colon), healthy humans skin and nares, eye; livestock (cows) and companion animal (dog),

hospital environment from central venous catheters, various public settings and from the natural environment, including from plants. In addition, a pangenome was constructed to identify the core and accessory genes in the genome. It was compared Clusters of Orthologous Groups (COG) function family of the unique genes found in *mecA*$^{+}$ *S. haemolyticus* isolates collected from public settings in East and West London, general public settings and in public areas in hospitals.

5.3 Results

5.3.1 Identification of virulence genes in *S. haemolyticus* isolates recovered from public settings in East and West London

11 virulence genes were identified in the 10 *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings from East and West London (Table 5.1). These included: the *ebp* gene (encoding elastin binding protein involved in bacterial adhesion); the *lip* gene (encoding lipase enzyme which is involved in the detachment of bacterial cells from colonised sites)*;* the *atl* gene (encoding autolysin involved in bacterial adhesion); capsule genes (involved in immune invasion); the *capB* and *capC* genes (encoding polyglutamic acid capsule for immune invasion), the *nuc* gene (encoding thurmonuclease an enzyme that can hydrolyse the host cell DNA and RNA) and the *cylR2* gene (encoding cytolysin, which is involved in lysing erythrocytes, polymorphonuclear leukocytes and macrophages). In addition, the *wbtP* gene (encoding for lipopolysaccharide modification in *Francisella* spp. essential for immune invasion) and the *sdrC* gene (encoding Ser-Asp rich fibrinogen-binding proteins which are involved in bacterial adhesion) were found in 20% of the isolates and the *clfB* gene (encoding clumping factor B which is involved in adhesion) were found in 10% of the isolates. Interestingly, all isolates had two capsule genes except for isolates 445 and 538, which had 14 capsule genes. It was also identified that there were two copies *cylR2* genes in sample 93, 373, 445 and 492. A hierarchical

cluster heatmap of public setting isolates shows there was no clustering of isolates based on location i.e. if they were recovered from West London or East London or general public settings or public areas in hospitals for virulent gene profiles (Figure 5.1).

Virulence genes	**Function**	**Percentage (%) of the *mecA* isolates from public areas East and West London**
nuc	Thermonuclease	100
lip	Lipase enzyme	100
capsule	Capsule genes	100
atl	Autolysin	100
ebp	Elastin binding protein	100
capB	Polyglutamic acid capsule	100
capC	Polyglutamic acid capsule	100
cylR2	Cytolysin enzyme	100
wbtp	lipopolysaccharide modification	20.0
sdrC	Ser-Asp rich fibrinogen-binding proteins	20.0
clfB	clumping factor B	10.0

Table 5.1: Percentage of Virulent genes found in the 10 *mecA*$^{+}$ *S. haemolyticus* from East and West London

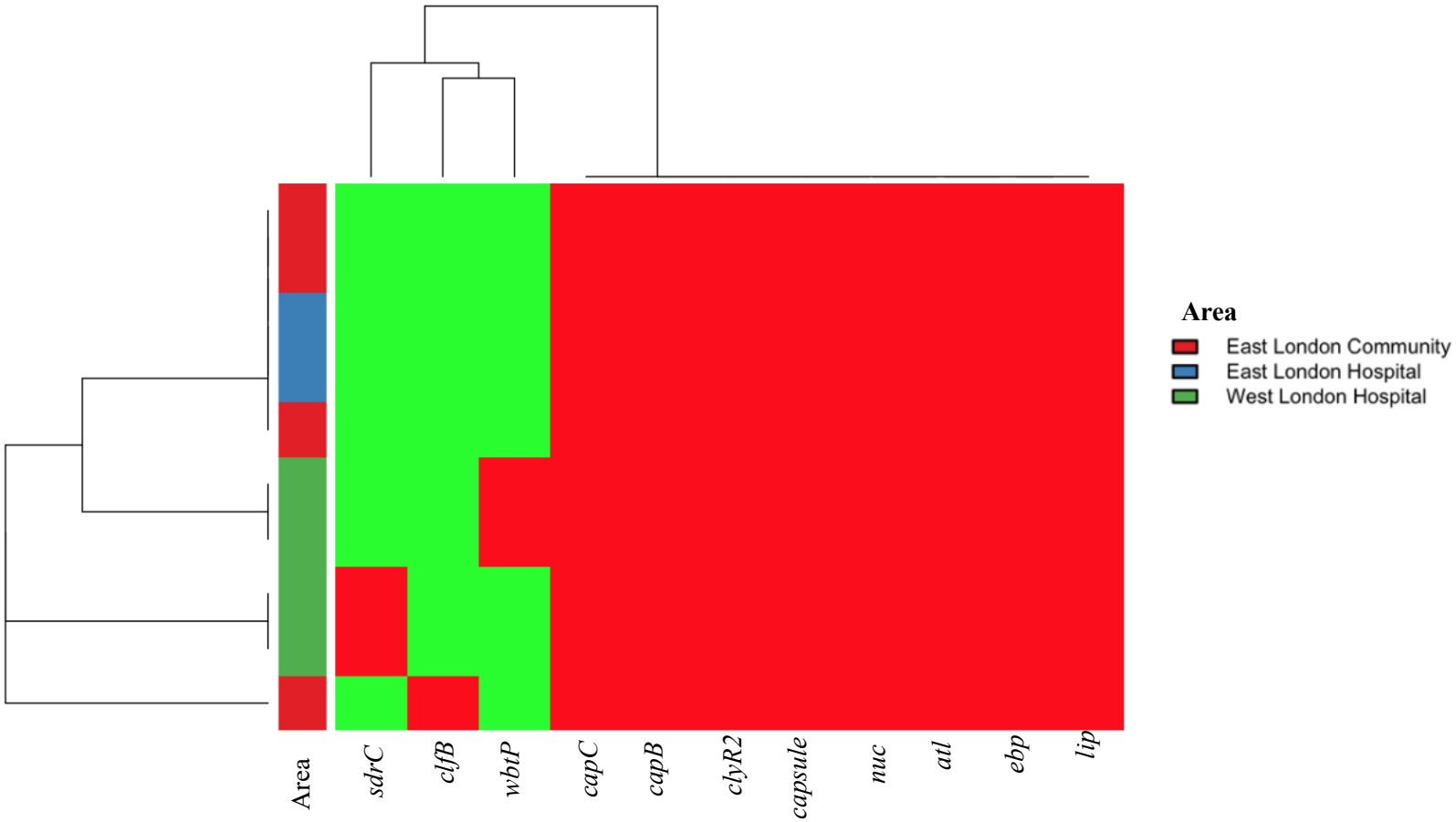


Figure 5.1: Hierarchy clustering heatmap of virulence genes found in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public setting in East and West London. Red tile - the presence of the virulent gene; green tile - the absence of the virulent gene**.**

5.3.2 Horizontal gene transfer events identified in whole genome sequenced *S. haemolyticus* isolates recovered from public settings from East and West London

Horizontal Gene Transfer (HGT) was detected in the 10 *mecA*$^{+}$ *S. haemolyticus* isolates from public settings in East and West London using the HGTector pipeline (Table 5.2). It was predicted that 8.9% to 11.2% of the isolate's genome were HGT-derived genes in comparison to the reference isolate JCSC 1435 recovered from humans skin in which 13.5% of its genome was predicted to be horizontally transferred based on BLAST hit distribution patterns from NCBI non-redundant protein sequences database (Pruitt et al., 2007; Zhu et al., 2014). One hundred and four genera were predicted to be donors for the *mecA*$^{+}$ *S. haemolyticus* isolates from those examined in this study based on the best hit of the non-redundant protein sequences database and NCBI taxonomy database. The highest portion of these genes predicted to be donated from *Bacillus* (mean n=31)*; Salinicoccus* (mean n=22) and *Micrococcus* (mean n=21) genera (Table 5.3).

***S. haemolyticus* ID**	**Number of protein-coding genes**	**Number of predicted HGT-derived genes**	**percentage of HGT derived genes**
JCSC 1435	2555	345	13.50
1	2344	221	9.43
93	2370	223	9.41
99	2575	228	8.85
105	2349	264	11.24
361	2418	224	9.26
373	2290	217	9.48
445	2403	264	9.32
492	2295	214	9.32
506	2380	217	9.12
538	2399	252	10.50

Table 5.2: Number of HGT genes in *mecA*$^{+}$ *S. haemolyticus* isolates from East and West London and reference *S. haemolyticus* isolate JCSC 1435

	S. haemolyticus ID										
Predicted Donor genus	**JCSC 1435**	**1**	**93**	**99**	**105**	**361**	**373**	**445**	**492**	**506**	**538**
Aeribacillus	0	1	1	1	0	1	0	2	1	1	1
Aerococcus	2	1	0	1	1	1	0	1	2	2	1
Alicyclobacillus	1	1	1	1	1	2	1	2	1	1	1
Alkalibacter	1	0	0	0	0	0	0	1	0	0	1
Alkalibacterium	0	0	0	0	0	1	0	0	0	0	1
Amphibacillus	0	0	0	0	0	0	0	0	2	0	0
Anaerovorax	0	1	0	0	1	0	0	0	0	0	0
Aneurinibacillus	2	2	2	2	2	2	2	2	2	2	2
Anoxybacillus	1	0	1	0	0	1	1	1	0	0	0
Anthococcus	0	1	1	0	1	0	0	0	1	0	0
Atopobacter	1	1	1	1	1	1	1	1	1	1	1
Auricoccucs	3	9	6	7	8	4	3	3	6	3	4
Avibacterium	1	1	1	1	1	1	1	1	0	1	1
Bacillus	36	26	32	31	29	32	31	40	27	34	32
Blautia	1	1	0	2	1	1	0	1	0	1	0
Brevibacillus	6	5	6	6	5	6	7	5	4	6	6
Caenibacillus	1	0	0	0	0	0	0	0	1	0	0
Caloramator	1	0	0	0	0	0	0	1	0	0	1
Carnobacterium	7	4	5	5	4	5	5	7	4	6	7
Caryophanon	0	0	0	0	0	0	0	0	1	0	0
Clostridioides	0	0	0	0	0	0	0	0	0	1	0
Clostridium	8	4	4	4	4	4	4	7	6	4	6
Corynebacterium	8	1	3	2	1	0	1	1	1	3	2
Cupriavidus	0	0	0	0	0	0	0	0	0	0	1
Cutibacterium	1	1	1	1	1	3	1	4	2	4	3
Desulfitobacterium	1	0	0	0	0	0	0	0	0	0	0
Domibacillus	2	1	2	1	1	1	1	0	1	0	0
Effusibacillus	0	0	0	0	0	0	0	0	1	0	0
Enterobacter	0	0	0	0	0	0	1	0	0	0	1
Enterococcus	34	8	9	7	8	6	9	9	11	12	7
Eremococcus	0	0	0	0	0	0	0	1	0	0	0
Erysipelatoclostridium	1	1	0	1	1	1	0	1	0	0	0
Eubacterium	0	0	0	0	0	0	0	1	0	0	0
Exiguobacterium	0	1	0	1	1	2	0	2	2	1	2
Faecalicatena	0	0	1	1	0	0	0	0	0	0	0
Faecalitalea	1	0	0	1	0	0	0	1	0	0	0
Fructobacillus	0	0	0	0	0	0	0	0	0	0	1
Gemella	1	0	0	0	0	0	0	0	0	0	0
Gemmiger	1	0	1	1	0	1	1	1	0	1	0
Geobacillus	2	0	0	0	1	0	1	1	0	0	1
Globicatella	0	0	0	1	0	0	0	0	0	1	0
Gracilibacillus	1	1	1	1	1	1	1	3	1	3	3
Haemophilus	1	0	1	1	0	1	1	1	1	1	1
Halarchaeum	0	0	1	0	0	1	1	1	0	1	1
Halobacillus	1	0	0	3	0	1	0	0	2	0	0
Halobacterium	0	1	0	1	1	0	0	0	0	0	0
Halolactibacillus	0	0	0	1	0	0	0	0	0	0	0

	S. haemolyticus ID										
Predicted Donor genus	**JCSC 1435**	**1**	**93**	**99**	**105**	**361**	**373**	**445**	**492**	**506**	**538**
Helcococcus	1	0	0	0	0	0	0	1	0	0	1
Jeotgalibaca	1	1	1	1	1	2	1	1	1	1	1
Jeotgalibacillus	1	0	0	0	0	0	0	0	0	0	0
Jeotgalicoccus	14	16	13	17	16	16	13	11	12	9	11
Kurthia	4	3	2	1	3	3	3	2	2	1	3
Kutzneria	0	0	0	1	0	0	0	0	0	0	0
Lacimicrobium	1	0	0	0	0	0	0	1	0	0	1
Lactobacillus	7	8	8	10	8	8	10	9	2	7	8
Lactococcus	0	0	1	1	0	0	1	1	0	0	0
Leptotrichia	1	1	1	1	1	1	1	1	2	1	1
Listeria	8	6	6	4	6	6	6	8	6	5	8
Loktanella	0	0	0	0	0	1	0	0	0	0	0
Lysinibacillus	11	8	12	9	8	9	11	9	8	9	11
Macrococcus	25	19	21	30	20	19	23	27	19	17	27
Maribacter	0	0	0	0	0	1	0	0	0	0	0
Marinilactibacillus	1	1	1	1	1	1	1	1	0	1	1
Marinococcus	0	0	0	1	0	0	0	0	0	1	1
Mycoplasma	0	0	0	0	0	2	0	0	0	0	0
Nosocomiicoccus	8	2	2	2	2	3	3	7	5	0	7
Novibacillus	0	1	0	0	0	1	1	1	1	0	1
Oceanobacillus	6	10	9	13	10	10	9	8	6	8	6
Ochrobactrum	1	1	1	2	1	2	1	1	2	1	1
Oenococcus	0	1	0	0	1	0	0	0	0	0	0
Oscillibacter	0	0	0	1	0	0	0	0	0	1	0
Paenibacillus	8	9	6	9	7	8	7	10	6	10	10
Paucisalibacillus	1	0	1	1	1	0	0	0	0	1	0
Pediococcus	1	0	0	0	0	1	0	2	2	1	0
Peptoclostridium	1	1	1	1	1	1	1	1	1	1	1
Peptoniphilus	1	0	0	0	0	1	0	1	0	1	1
Planococcus	1	2	2	3	2	4	3	1	2	3	1
Planomicrobium	1	2	2	1	1	1	1	1	2	2	0
Pseudomonas	23	1	1	2	1	1	1	2	1	0	4
Rheinheimera	1	1	1	1	1	1	1	1	0	1	1
Rhizobium	1	1	0	1	1	1	0	2	1	2	1
Rhodococcus	0	1	0	1	1	1	0	0	1	1	1
Saccharomonospora	0	0	1	0	0	0	1	1	0	0	0
Salinicoccus	35	23	21	24	25	17	18	23	24	20	24
Salsuginibacillus	1	0	0	0	1	0	0	0	1	1	0
Selenomonas	0	1	1	1	1	0	1	0	0	0	1
Sharpea	0	0	1	0	0	0	1	0	0	0	0
Solibacillus	0	0	0	0	0	0	0	0	1	0	0
Sporosarcina	0	0	1	1	1	1	0	0	3	0	0
Streptococcus	11	9	9	11	9	9	9	9	9	8	10
Streptomyces	1	2	0	2	2	0	0	0	0	3	0
Terribacillus	1	2	2	1	2	0	2	1	1	0	1
Tetragenococcus	4	1	1	1	1	2	2	2	1	2	4
Thalassobacillus	0	0	0	0	0	0	0	0	1	0	0

	S. haemolyticus ID										
Predicted Donor genus	**JCSC 1435**	**1**	**93**	**99**	**105**	**361**	**373**	**445**	**492**	**506**	**538**
Thermoactinomyces	2	0	2	0	0	2	2	2	0	0	0
Tissierella	0	1	0	0	1	0	0	0	0	0	0
Tolumonas	1	1	0	0	1	0	0	0	0	1	0
Trichococcus	2	0	1	1	0	1	1	2	0	1	3
Tuberibacillus	0	0	0	0	0	0	0	0	1	0	0
Vagococcus	1	2	1	3	2	3	1	1	1	0	0
Veillonella	2	2	0	0	2	1	0	3	2	0	0
Virgibacillus	2	2	1	0	2	0	1	1	2	1	4
Viridibacillus	1	0	1	0	0	1	1	1	1	1	1
Weissella	2	3	3	2	3	3	3	3	2	2	3

Table 5.3: *MecA*$^{+}$ *S. haemolyticus* and reference JCSC 1435 HGT-derived from by the best match putative donor genus as indicated by the best distal match.

13 out of 14 antibiotic resistance genes in *mecA*$^{+}$ *S. haemolyticus* isolates in this study were predicted to be horizontally transferred (Table 5.4). The only gene not predicted to be horizontally transferred was *blaZ*. Antibiotic resistance gene *AAC(6')-Ie-APH(2'')-Ia* and *APH(3')-IIIa* were predicted to be donated by a different organism in different isolates. *AAC(6')-Ie-APH(2'')-Ia* was predicted to be donated by 3 species in isolate 1, 93, 99, 105. These species were *Clostridiales bacterium VE202-16*, *Streptococcus mitis B6* and *Enterococcus faecalis V583*. *APH(3')-IIIa* was predicted to be donated by 2 species in isolate 93 and 99. These species were *Streptococcus mitis B6* and *Enterococcus faecium DO*.

S. haemolyticus ID.	Antibiotic resistance genes	Predicted donor organism
1	*ANT(4')-Ib* *MecA* *lnuA* *AAC(6')-Ie-APH(2'')-Ia* *tet(k)* *dfrG*	*Arcobacter thereius* *Macrococcus canis* *Lactobacillus johnsonii* *Clostridiales bacterium VE202-16* *Streptomyces cinnamoneus* *Vagococcus teuberi*
93	*AAC(6')-Ie-APH(2'')-Ia* *mecA* *dfrG* *APH(3')-IIIa*	*Streptococcus mitis B6* *Macrococcus canis* *Vagococcus teuberi* *Streptococcus mitis B6*
99	*lnuA* *AAC(6')-Ie-APH(2'')-Ia* *mecA* *dfrG* *APH(3')-IIIa* *tet(k)*	*Lactobacillus johnsonii* *Enterococcus faecalis V583* *Macrococcus canis* *Vagococcus teuberi* *Enterococcus faecium DO* *Streptomyces cinnamoneus*
105	*ANT(4')-Ib* *lnuA* *AAC(6')-Ie-APH(2'')-Ia* *mecA* *tet((k)* *dfrG*	*Arcobacter thereius* *Lactobacillus johnsonii* *Clostridiales bacterium VE202-16* *Macrococcus canis* *Streptomyces cinnamoneus* *Vagococcus teuberi*
361	*fusB* *mecA* *qacA/B*	*Enterococcus thailandicus* *Macrococcus canis* *Bacillus ndiopicus*
373	*dfrG* *mecA* *fusB*	*Vagococcus teuberi* *Macrococcus canis* *Enterococcus thailandicus*
445	*mphC* *msrA* *mecA* *vgaA* *qacA/B*	*Veillonella atypica* *Veillonella atypica* *Macrococcus canis* *Aeribacillus pallidus* *Bacillus ndiopicus*
492	*mphC* *mecA* *msrA*	*Veillonella atypica* *Macrococcus canis* *Veillonella atypica*
506	*qacA/B* *mecA* *ermA* *tet(K)*	*Bacillus ndiopicus* *Macrococcus canis* *Enterococcus* spp. *HMSC29A04* *Streptomyces cinnamoneus*
538	*qacA/B* *mecA* *fusB*	*Bacillus ndiopicus* *Macrococcus canis* *Enterococcus thailandicus*

Table 5.4: ***MecA*$^{+}$ *S. haemolyticus*** **isolates from public settings in East and West London and reference *S. haemolyticus* isolate JCSC 1435 HGT-derived genes by their putative donor genus as indicated by the best distal match.**

Only 3 virulence genes were predicted to be horizontally transferred (Table 5.5). Interestingly the *cylR2* gene, which was found in all the isolates were also predicted to be horizontally transferred. 9 of the 14 predicted capsule genes in isolates 445 and 538 were predicted to be horizontally transferred. 4 of the 9 capsule genes that were horizontally transferred were predicted to be donated by *Nosocomiicoccus* genus, whereas the other 3 capsules genes were predicted to be donated by *Listeria grayi* DSM 20601 and the other by *Bacillus massilionigeriensis and Caloramator mitchellensis*.

S. haemolyticus ID	genes	Predicted donor species
1	*cylR2*	*Virgibacillus* spp. *SK37*
	cylR2	*Streptococcus* spp. *HMSC10A01*
93	*cylR2*	*Streptococcus* spp. *HMSC10A01*
99	*cylR2*	*Streptococcus* spp. *HMSC10A01*
105	*cylR2*	*Virgibacillus* spp. *SK37*
	cylR2	*Streptococcus* spp. *HMSC10A01*
361	*cylR2*	*Streptococcus* spp. *HMSC10A01*
373	*cylR2*	*Streptococcus* spp. *HMSC10A01*
445	*Capsule*	*Nosocomiicoccus* spp. *HMSC067E10*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Nosocomiicoccus* spp. *HMSC059G07*
	Capsule	*Nosocomiicoccus ampullae*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Caloramator mitchellensis*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Bacillus massilionigeriensis*
	Capsule	*Nosocomiicoccus* spp. *HMSC09A07*
	CylR2	*Streptococcus* spp. *HMSC10A01*
492	*cylR2*	*Streptococcus gordonii str. Challis substr. CH1)*
506	*cylR2*	*Streptococcus* spp. *2_1_36FAA*
538	*wbtP*	*Lacimicrobium* spp. *SS2-24*
	cylR2	*Streptococcus* spp. *HMSC10A01*
	cylR2	*Virgibacillus* spp. *SK37*
	Capsule	*Nosocomiicoccus* spp. *HMSC09A07*
	Capsule	*Bacillus massilionigeriensis*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Caloramator mitchellensis*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Nosocomiicoccus ampullae*
	Capsule	*Nosocomiicoccus* spp. *HMSC059G07*
	Capsule	*Listeria grayi DSM 20601*
	Capsule	*Nosocomiicoccus* spp. *HMSC067E10*

Table 5.5: **Horizontally transferred virulence genes in *mecA*$^+$ *S. haemolyticus* isolates recovered from public areas from East and West London**

5.3.3 Antibiotic resistance genes carried within a genomic island *in mecA*$^+$ *S. haemolyticus* from public settings in East and West London

Genomic islands were predicted in the 10 *mecA*$^+$ *S. haemolyticus* isolates from East and West London based on the Islandviewer 4 method (Table 5.6). All samples were predicted to have regions in their genome that were considered as genomic islands. Interestingly,

the *mecA* gene was not predicted within a genomic island via this method but the *mecA* gene was found in near a predicted genomic island. Out of this 10 *S. haemolyticus* isolates 2 (20%) had 4 antibiotic resistance genes; 5 (50%) had 2 antibiotic resistance genes and 2 (20%) had 1 antibiotic resistance genes within genomic island regions. The most common antibiotic-resistance genes found within genomic island regions were *blaZ* (n=6); followed by *qacA/B* (n=3); *tet(K) (n=2); ANT(4')-IB*, *APH(3')-IIIa*, *fusB*, *vgaA*, *msrA*, *mphC* and *ermA* (n=1). No virulence genes were found within genomic island regions.

S. haemolyticus ID	No. of predicted genomic island regions	Antibiotic resistance genes
1	7	*lnuA* *AAC(6')-Ie-APH(2'')-Ia* *blaZ* *ANT(4')-IB*
93	4	*blaZ* *APH(3')-IIIa*
99	13	*blaZ* *AAC(6')-Ie-APH(2'')-Ia* *APH(3')-IIIa* *lnuA*
105	8	*tetK* *AAC(6')-Ie-APH(2'')-Ia*
361	6	*blaZ* *qacA/B*
373	7	*fusB*
445	7	*vgaA*
492	8	*msrA* *mphC*
506	6	*ermA* *qacA/B* *tet(k)* *blaZ*
538	5	*qacA/B* *blaZ*

Table 5.6: **Genomic islands of 10 *mecA*$^+$ *S. haemolyticus* isolates recovered from East and West London**

5.3.4 Antibiotic resistance and virulence genes carried on a plasmid.

All *mecA*$^+$ *S. haemolyticus* isolates from the public settings (n=10) were identified to carry between 1 to 5 plasmids. Three isolates had the majority of their antibiotic resistance genes encoded on the plasmids (Table 5.7). 8 out of 10 of these isolates carried antibiotic resistance genes of which 3 isolates (30%) had 5 antibiotic resistance genes carried on their plasmids; 1 isolate (10%) had 3 antibiotic resistance genes carried on their plasmids; 2 isolates (20%) had 2 antibiotic resistance genes carried on their plasmid and 1 isolate (10%) had 1 antibiotic resistance gene carried on their plasmid. Interestingly, isolate 445 had 5 of its antibiotic resistance genes on a single plasmid. *BlaZ* (n=6) was the most

common antibiotic resistance gene found on the plasmid; followed by *AAC(6')-Ie-APH(2'')-Ia*, *qacA/B* and *lnuA* (n=3)*; tet(K)* (n=2) and *mphC*, *msrA* and *vgaA* (n=1). There were no detectable virulence genes to be encoded on any of the *mecA*$^{+}$ *S. haemolyticus* isolates from the public settings plasmids.

***S. haemolyticus* ID**	**Number of Plasmids**	**Chromosome: plasmid-encoded antibiotic genes**	**Antibiotic resistance genes (and plasmids they were found on)**
1	2	2:5	*AAC(6')-Ie-APH(2'')-Ia* (plasmid 1) *blaZ* (plasmid 1) *tet(K)* (plasmid 2) *ANT(4')-Ib* (plasmid 2) *lnuA* (plasmid 2)
93	3	5:0	No genes
99	5	4:3	*lnuA* (plasmid 1) *blaZ* (plasmid 1) *AAC(6')-Ie-APH(2'')-Ia* (plasmid 2)
105	2	2:5	*AAC(6')-Ie-APH(2'')-Ia* (plasmid 1) *blaZ* (plasmid 1) *tet(K)* (plasmid 2) *ANT(4')-Ib* (plasmid 2) *lnuA* (plasmid 2)
361	3	2:2	*blaZ* (plasmid 1) *qacA/B* (plasmid 1)
373	1	4:0	No genes
445	2	1:5	*qacA/B* (plasmid 1) *mphC* (plasmid 1) *msrA* (plasmid 1) *blaZ* (plasmid 1) *vgaA* (plasmid 1)
492	2	3:0	
506	3	4:1	*qacA/B* (plasmid 1)
538	1	1:2	*qacA/B* (plasmid 1) *blaZ* (plasmid 1)

Table 5.7: **Number of plasmids and the antibiotic resistance genes carried on plasmids identified in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from East and West London.**

5.3.5 Phage prediction

Phage insertion was identified in all the *mecA*$^{+}$ *S. haemolytic*us isolates recovered from public setting to different degrees of completeness using the PHASTER software (Table 5.8). 6 of the isolates were identified to have incomplete phage insertion, whereas 3

isolates (93, 99 and 361) had intact phage sequences. The intact phages insertion was identified as known 'phages to infect *Staphylococcus*'. Two phages insertion sequences were identified with proteins showing homology to phage proteins associated with infecting *Synechococcus* and *Prochlorococcus* genus. Isolate 445 had phage inserted into 6 regions in the genome, isolate 492 had phage insertion in 5 regions in the genome. Isolate 99 was identified to have phage insertion in 3 regions in the genome, whereas the other 8 isolates (80%) had phage insertion in 1 region of their genome. *APH(3')-IIIa* and *fusB* antibiotic resistance genes were identified to be within the phage insertion region for isolate 93 and isolate 538, respectively.

.

S. haemolyticus ID	Most common phages	Completeness	Antibiotic resistance gene
1	*Staphylococcus phage PT1028*	Incomplete	
93	*Staphylococcus phage CNPx*	Intact	*APH(3')-IIIa*
99	*Staphylococcus phage PT1028*	Incomplete	
	Staphylococcus phage StB27	Intact	
	Staphylococcus phage StB12	Intact	
105	*Staphylococcus phage PT1028*	Incomplete	
361	*Staphylococcus phage StB12*	Intact	
373	*Staphylococcus phage PT1028*	Incomplete	
445	*Synechococcus phage S-SSM7*	Incomplete	
	Staphylococcus phage PT1028	Incomplete	
	Staphylococcus phage PT1028	Incomplete	
	Staphylococcus phage StB2- like	Incomplete	
	Staphylococcus phage CNPH82	Incomplete	
	Staphylococcus phage Spbeta-like	Incomplete	
492	*Staphylococcus phage Spbeta-like*	Incomplete	
	Staphylococcus phage PT1028	Incomplete	
	Staphylococcus phage CNPH82	Incomplete	
	Staphylococcus phage PT1028	Incomplete	
	Prochlorococcus phage P-SSM7	Incomplete	
506	*Staphylococcus phage PT1028*	Incomplete	
538	*Staphylococcus phage PT1028*	Incomplete	*fusB*

Table 5.8: **Phage prediction in *mecA*$^+$ *S. haemolyticus* isolates recovered from East and West London; identification of antibiotic resistance genes carried by phages**

5.3.6 Phylogenetic analyses of *mecA*$^+$ *S. haemolyticus* isolates recovered from East and West London compared with S. *haemolyticus* reference isolates from the ENA database

Phylogenetic analyses were performed for the 10 *mecA*$^+$ *S. haemolyticus* isolates recovered from public areas from East and West London to compare their relatedness with reference and draft isolates from the ENA database that was recovered from different sources. These sources included isolates recovered from clinical samples (n=48); human commensal isolates from skin, nares and eyes (n=7); livestock (cow) and companion animal (dog) isolates (n=8); other public settings and natural environment isolates (n=5), hospital environment from catheters from central venous catheter (n=2) and plant-associated isolates (n=3). From the core SNP phylogenetic tree, two distinctive clades

were identified (Figure 5.2). Clade A consists of isolates from the ENA database recovered from clinical samples (blood, vagina and sputum), livestock (cows), a companion animal (dog), groundwater and a healthy human eye. Clade B consist of isolates from the ENA database recovered from clinical samples (blood, eye and colon) healthy human skin, central venous catheter, Kefir seed, willow tree, livestock (cows), tropical air samples, copper alloy coin, surface area of a building and waste and hygiene compartment of the International Space Station. All *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public areas in London were found in clade B, but one isolate (492) recovered from West London hospital belonged to clade A. Genetic relatedness between isolates recovered in this study from public areas in hospital (373, 445 and 538) and isolates (1, 93, 99 and 105) recovered from public setting to that of the clinical isolates from the ENA database recovered from an eye (SH1572), blood (M-176), and central venous catheter (95671). Interestingly, one isolate recovered from public areas in hospitals (492) in this study were genetically related to an isolate from the ENA database recovered from a dog (SW007); and another isolate recovered from public areas in hospitals (445) was genetically similar to ENA database isolate from Kefir seeds (OG2). Additionally, from the ENA database isolates two livestock-associated isolates from cows (BC05211 and NW19) and two isolates recovered from a plant (OG2 from Kefir seeds and RIT283 from willow) to be genetically related to isolates recovered from clinical blood samples from the ENA database. There were no *S. haemolyticus* isolates from the ENA database which belong to the same genetic lineage of isolate 506 which was recovered in public areas in the hospital in West London.

.

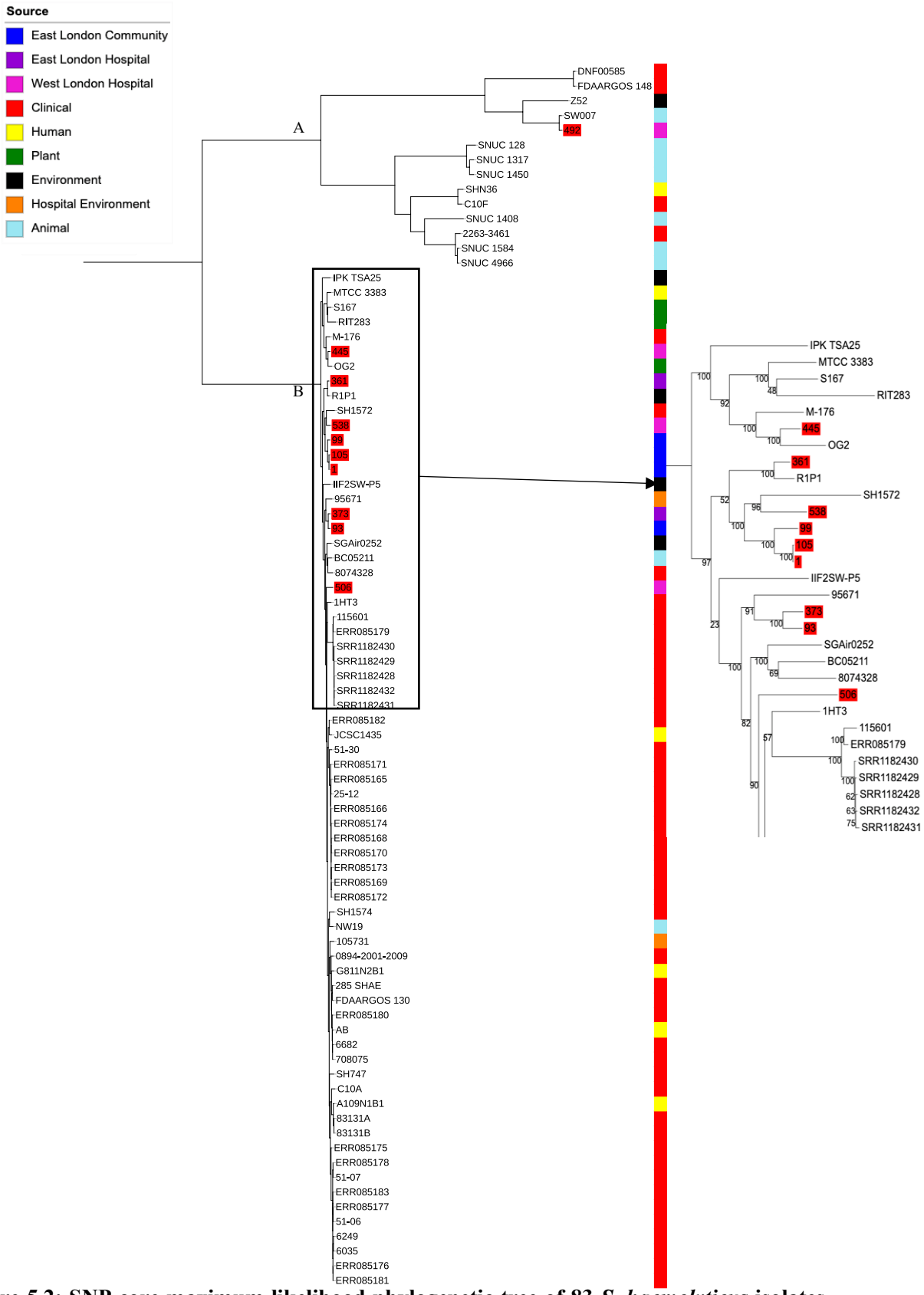


Figure 5.2: SNP core maximum likelihood phylogenetic tree of 83 *S. haemolyticus* isolates recovered from different sources and *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings in East and West London. Red background labels *mecA*$^{+}$ *S. haemolyticus* isolates from public settings in East and West London.

PopPUNK analyses revealed that there were 38 combined core and accessory gene clusters predicted of which the accessory genome was found within 11 clusters (Figure 5.3). Five of these clusters had isolates that were identified to be from the same combined cluster and 6 clusters had isolates that only contained isolates from the ENA database that were recovered from clinical samples. Interestingly, isolates from the East London Community and the East London Hospital were found together in the same cluster despite not always having the same combined core and accessory cluster (Cluster 3). West London isolates can be found in multiple different clusters (Clusters 2, 3 and 6). The accessory genomes of all isolates recovered from East London had genetic similarities with isolates from the ENA database recovered from clinical samples from an eye (SH1572), venous catheter (95671) and environmental isolates recovered from a copper alloy coin (R1P1), whereas the accessory genomes of isolates recovered from West London were genetic similarities to isolates from ENA database recovered from clinical samples from an eye (SH1572 and SH1574), colon (1HT3), blood (FDAARGOS-148), vagina (DNF00585) and sputum (C10F), healthy humans, plants (RIT283 and S167) , livestock (NW19) and companion animal (SW007).

.

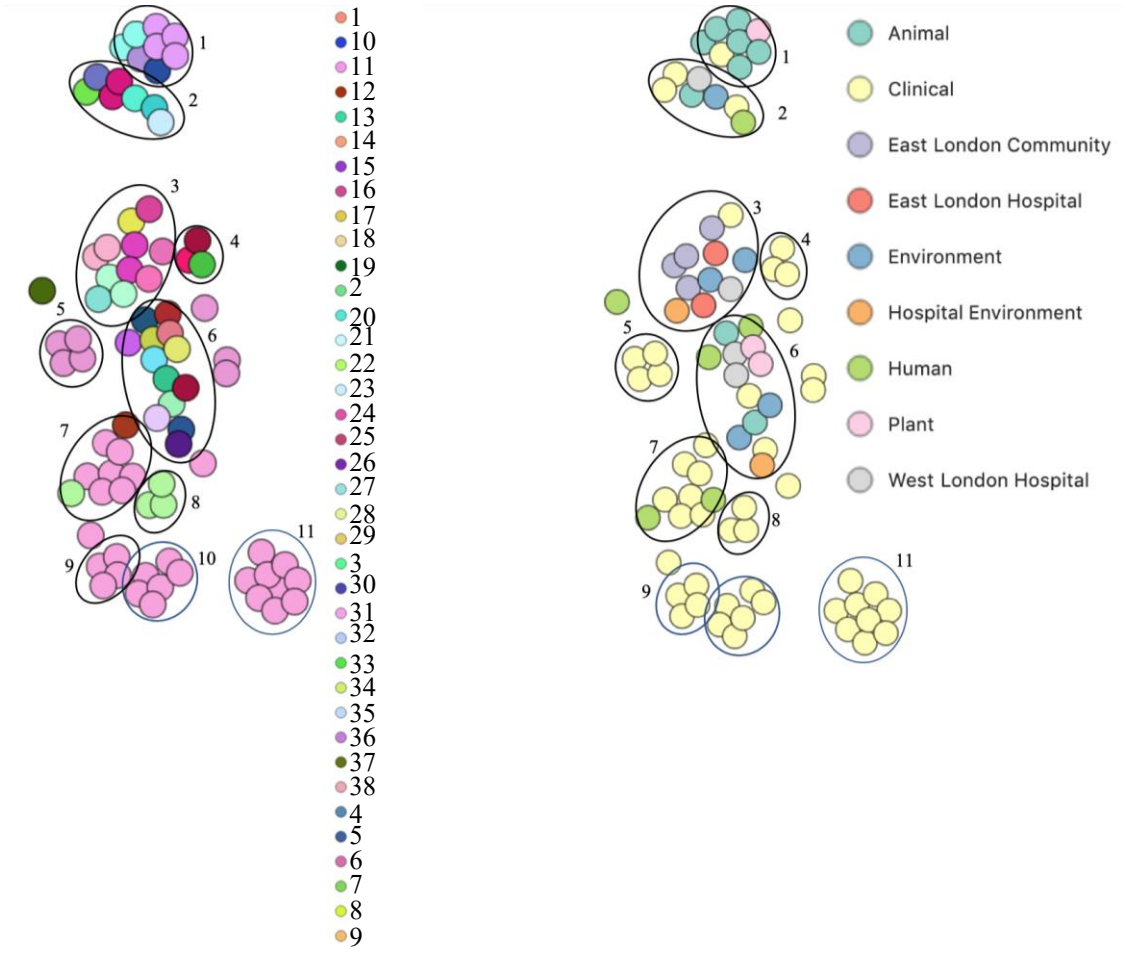


Figure 5.3: **t-SNE analyses of the distance of the accessory genome in 83 *S. haemolyticus* isolates including *mecA*$^{+}$ S. *haemolyti*cus isolates recovered from public settings in East and West London.** The Analysis was performed using PopPUNK pipeline Maximum number of mixture components was set at 4 and for the perplexity of the t-SNE set at 15. (A) Combined cluster from PopPUNK analyses; (B) isolation source.

5.3.7 Pangenome analysis of *S. haemolyticus*

A pangenome analysis was performed on the 83 *S. haemolyticus* genomes that were used in the phylogenetic analysis, including the 10 *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings in East and West London. In total, 8,978 genes were identified in the pangenome of which 1,200 (13.4%) were considered to be core genes. 458 (5.1%) of these genes were considered to be softcore genes (95 to 99% of isolates have these genes); 1,218 (13.6%) were considered to be shell genes (15 to 95% of isolates have these genes) and 6,102 (68.0%) were considered to be cloud genes (≤ 15% of isolates have these genes). Hierarchy clustering heatmap of all 8,978 genes found in the pangenome shows there was no clustering by their isolation source (Figure 5.4).

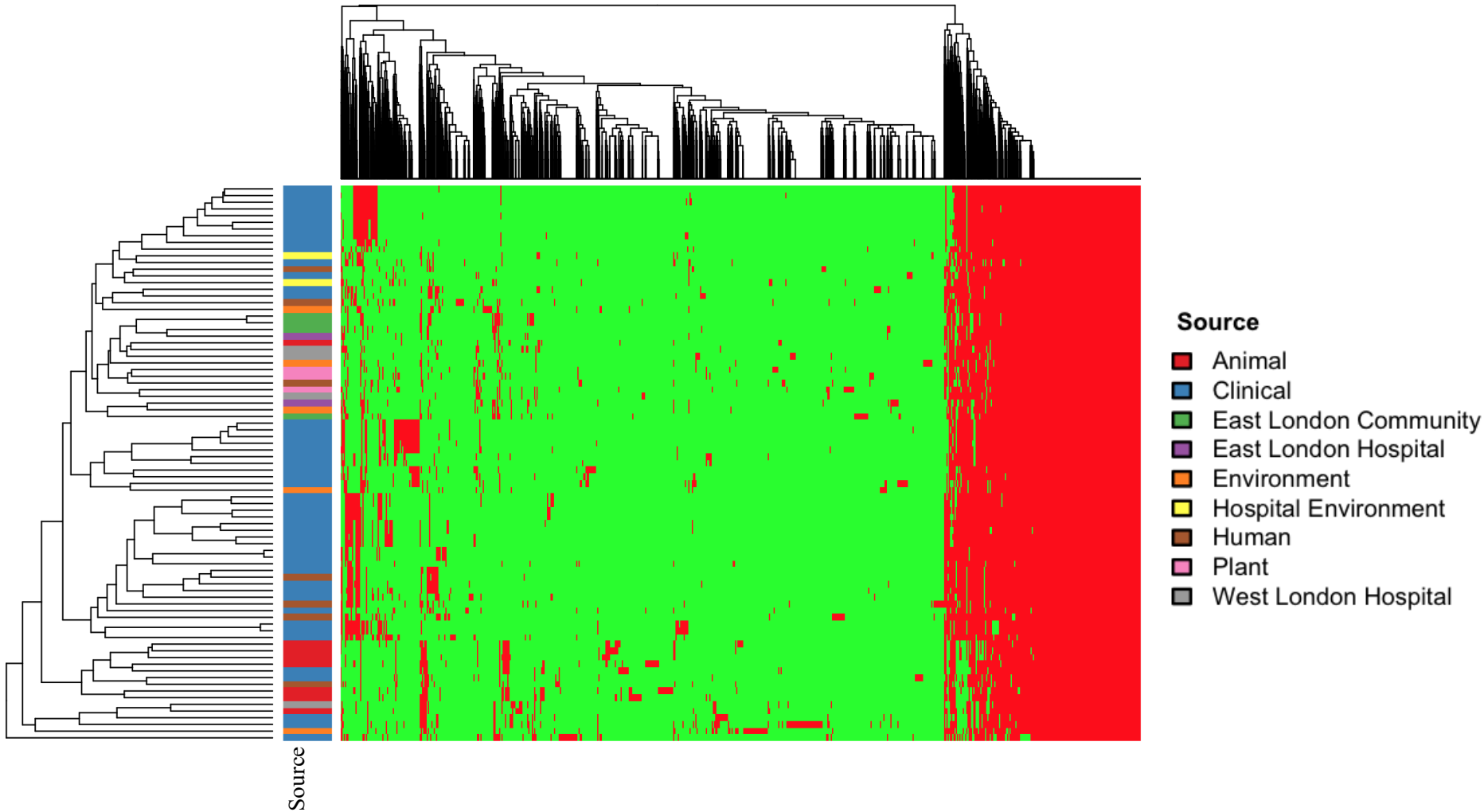


Figure 5.4: **Hierarchy clustering heatmap of 8,978 genes found in pangenome of 83 *S. heamolyticus* isolates base on their isolation source.** Red tile presence of the gene; green tile absence of the gene.

23 antibiotic resistance genes and 18 virulence genes were identified in the pangenome (Table 5.9 and 5.10). The *nuc* gene and capsule gene were ubiquitous in all 83 *S. haemolyticus* isolates. *MecA* and *qacA/B* were identified in 78.3 and 72.3% of *S. haemolyticus* isolates, respectively. *FosB3* and *dfrC* antibiotic resistance genes were unique to isolates from the ENA database recovered from clinical samples where *eno* virulence gene (*Streptococcal* enolase enzyme responsible for adhesion) was unique to a clinical isolate. Multiple capsule genes were identified in these isolates. 34 (41.0%) of the isolates had 2 capsule genes; 3 (3.6%) isolates had 3 capsule genes; 1 (1.2%) isolate had 4 capsule genes; 3 (3.6%) isolates had 6 capsule genes; 7 (8.4%) isolates had 7 capsule genes; 11 (13.3%) isolates had 14 capsule genes; 5 (6.0%) isolates had 15 capsule genes; 1 (1.2%) isolate had 16 capsule genes; 3 (3.6%) isolates had 17 capsule genes and 15 (18.1%) isolates had 18 capsule genes. The number of capsule genes was not unique to a single isolation source (Table 5.11).

A hierarchy clustered heatmap showed there was no clustering by isolation source for antibiotic resistance genes or virulence (Figure 5.5). It did, however, show that isolates from public settings had similar antibiotic resistance and virulence genes profiles of the isolates from the ENA database recovered from clinical samples, livestock, companion animal, healthy humans and other public settings and natural environments.

Antibiotic resistance genes	List of antibiotic classes	Percentage (%)
AAC(6')-Ie-APH(2'')-Ia	Aminoglycoside	65.1
dfrG	Diaminopyrimidine	25.3
mecA	Penam, monobactam, cephalosporin, carbapenem, cephamycin	78.3
blaZ	Penam	72.3
ermC	lincosamide, streptogramin, macrolide	9.6
qacA/B	Fluoroquinolone	56.6
mphC	Macrolide phosphotransferase	47.0
mupA	Mupirocion	10.8
msrA	Streptogramin, macrolide	51.8
vgaALC	Streptogramin, pleuromutilin	8.4
dfrC	Diaminopyrimidine	12.1
APH(3')-IIIa	Aminoglycoside	37.3
SAT-4A	Nucleoside antibiotic	32.5
fusC	Fusidic acid	9.6
tet(K)	Tetracycline	25.3
ermC	Lincosamide, streptogramin, macrolide	15.7
erm(33)	Macrolide, Streptogramin, lincosamide	2.4
fusB	Fusidic acid	21.7
cat	Phenicol	6.02
ANT(4')-Ib	Aminoglycoside	13.3
FosB3	Fosfomycin	4.8
lnuA	Lincosamide	4.8
ermA	Macrolide, Streptogramin, lincosamide	2.4

Table 5.9: **Percentage of antibiotic resistance genes present in 83 *S. haemolyticus* isolate in pangenome analysis**

Virulence genes	Gene function	Percentage (%)
atl	Autolysin	96.4
clfB	Clumping factor B	3.6
ebp	Elastin binding protein	98.8
lip	Lipase	97.6
nuc	Thermonuclease	100.0
capsule	capsule	100.0
capB	Polyglutamic acid capsule	86.8
cylR2	Cytolysin	95.2
capC	Polyglutamic acid capsule	86.8
polysaccharide capsule	polysaccharide capsule	7.2
wbtE	LPS	10.8
sdrC	Ser-asp fibrinogen-binding protein	6.0
sdrE	Ser-asp fibrinogen-binding protein	9.6
Capsule (Acinetobacter)	Capsule (*Acientobacter*)	4.8
uge	Capsule (*Klebsiella*)	6.0
eno	Streptococcal enolase	1.2
sdrD	Ser-asp fibrinogen-binding protein	3.6
wbtP	LPS	9.6

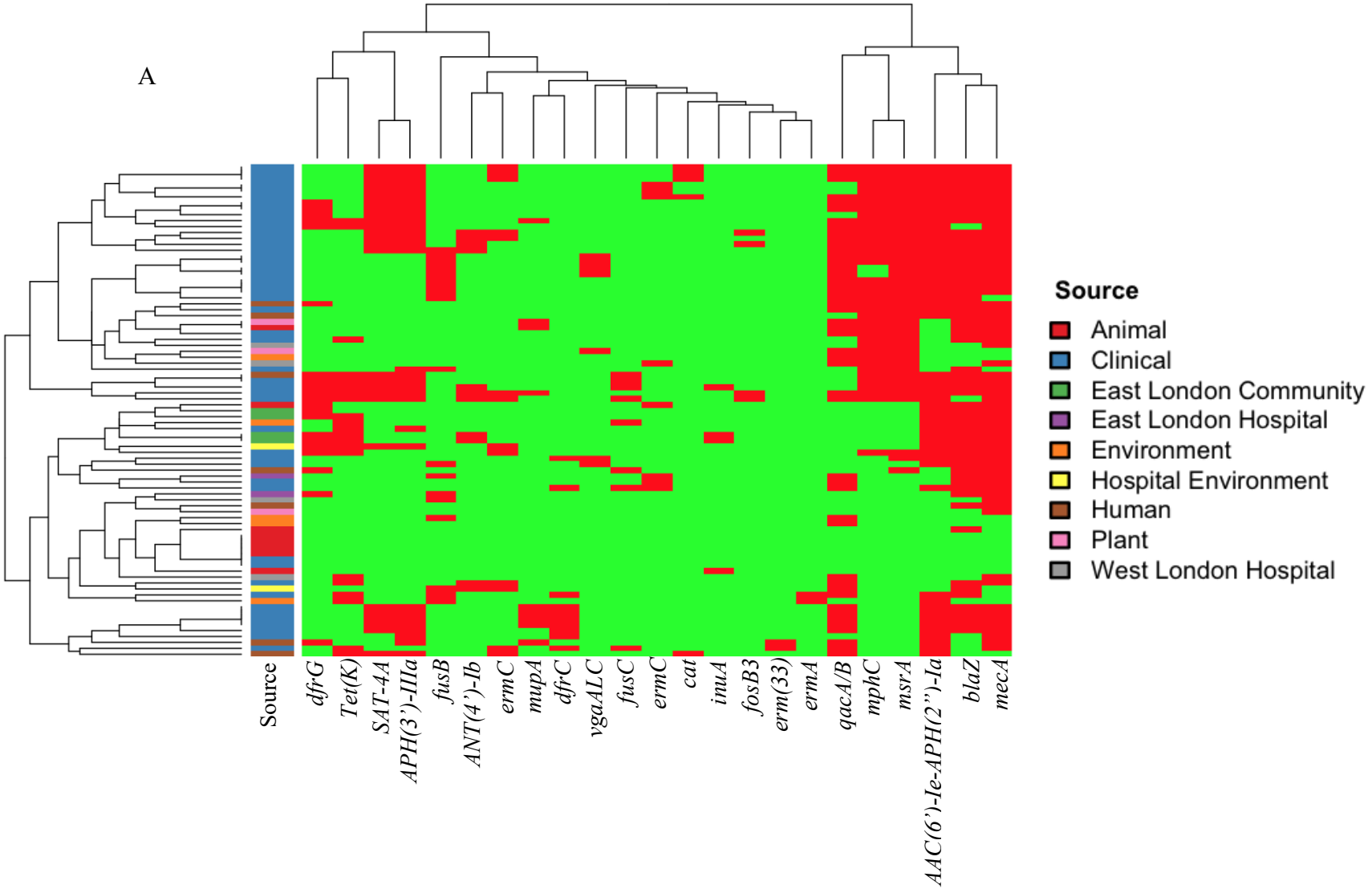
Table 5.10: Percentage of virulence genes present in 83 *S. haemolyticus* isolate in pangenome analysis

S. haemolyticus ID	Source	Number of capsule genes
BC05211	Animal	2
SW007	Animal	2
M-176	Clinical	2
1HT3	Clinical	2
25-12	Clinical	2
6035	Clinical	2
6249	Clinical	2
8074328	Clinical	2
95671	Hospital environment	2
ERR085165	Clinical	2
ERR085166	Clinical	2
ERR085168	Clinical	2
ERR085169	Clinical	2
ERR085170	Clinical	2
ERR085171	Clinical	2

S. haemolyticus ID	Source	Number of capsule genes
ERR085172	Clinical	2
ERR085173	Clinical	2
ERR085174	Clinical	2
ERR085176	Clinical	2
ERR085181	Clinical	2
FDAARGOS_148	Clinical	2
SH1574	Clinical	2
SH747	Clinical	2
1	East London Community	2
105	East London Community	2
93	East London Community	2
99	East London Community	2
361	East London Hospital	2
373	East London Hospital	2
IPK_TSA25	Environment	2
R1P1	Environment	2
105731	Hospital environment	2
492	West London Hospital	2
506	West London Hospital	2
SNUC_1450	Animal	3
DNF00585	Clinical	3
SH1752	Clinical	3
G811N2B1	Human	4
2263-3461	Clinical	6
C10F	Clinical	6
SHN36	Human	6
OG2	Plant	7
SNUC_128	Animal	7
SNUC_1408	Animal	7
C10A	Clinical	7
IIF2SW-P5	Environment	7
Z52	Environment	7
MTCC_3383	Human	7
0894-2001-2009	Clinical	14
285_SHAE	Clinical	14
6682	Clinical	14
708075	Clinical	14
ERR085175	Clinical	14
ERR085180	Clinical	14
FDAARGOS_130	Clinical	14
AB	Human	14
JCSC1435	Human	14
445	West London Hospital	14
538	West London Hospital	14

S. haemolyticus ID	Source	Number of capsule genes
51-06	Clinical	15
51-07	Clinical	15
ERR085177	Clinical	15
ERR085178	Clinical	15
NW19	Human	15
SNUC_1317	Animal	16
SNUC_4966	Animal	17
A109N1B1	Human	17
RIT283	Plant	17
SNUC_1584	Animal	18
115601	Clinical	18
51-30	Clinical	18
83131A	Clinical	18
83131B	Clinical	18
ERR085179	Clinical	18
ERR085182	Clinical	18
ERR085183	Clinical	18
S167	Plant	18
SRR1182428	Clinical	18
SRR1182429	Clinical	18
SRR1182430	Clinical	18
SRR1182431	Clinical	18
SRR1182432	Clinical	18
SGAir0252	Environment	18

Table 5.11: Number of capsule genes to the isolation sources found in the 83 *S. haemolyticus* isolates. Red highlight = East London Community, Blue label= East London Hospital, Green label= West London Hospital


A
Source
dfrG
Tet(K)
SAT-4A
APH(3')-IIIa
fusB
ANT(4')-Ib
ermC
mupA
dfrC
vgaALC
fusC
ermC
cat
inuA
fosB3
erm(33)
ermA
qacA/B
mphC
msrA
AAC(6')-Ie-APH(2'')-Ia
blaZ
mecA
Source
Animal
Clinical
East London Community
East London Hospital
Environment
Hospital Environment
Human
Plant
West London Hospital

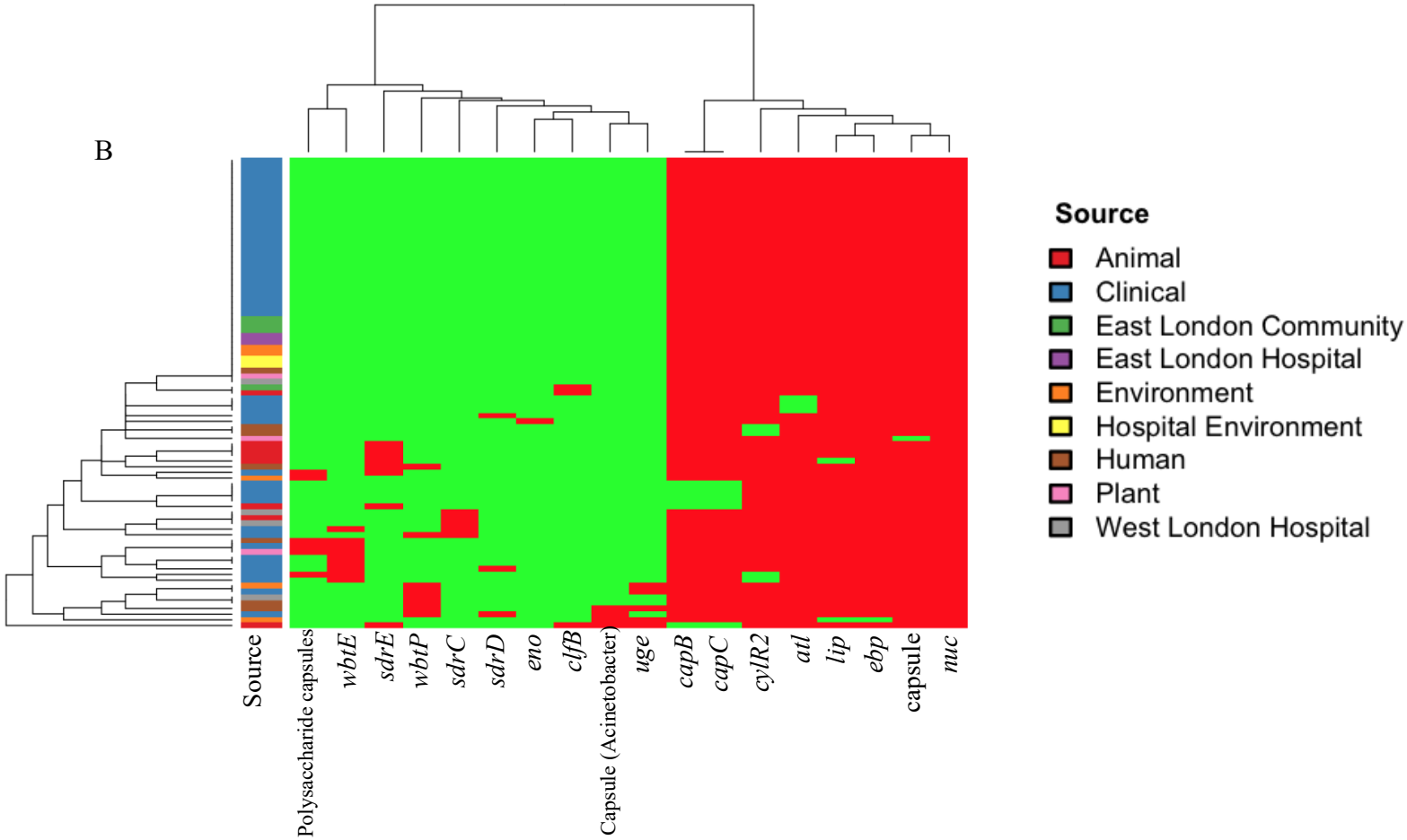


Figure 5.5: **Hierarchy cluster heatmap of antibiotic resistance and virulence genes found in 83 *S. haemolyticus* isolates used in the pangenome analysis**

(A) Hierarchy clustered heatmap of antibiotic resistance genes found by source; (B) Hierarchy clustered heatmap of virulence gene by source.

Red tile present; green tile absent.

The COG family group was identified for the core and the accessory genome in *S. haemolyticus* pangenome analysis (Table 5.12). From the core genome, the highest proportion of the genes were classified as general function prediction only or function unknown genes (11.9% and 9.7% respectively). The core genes group translational, ribosomal structure and biogenesis and amino acid transport and metabolism (9.5% and 9.5% respectively) were the highest proportion of known functional COG group. For the accessory genome, the highest portion of the genes belonged to the COG group were those responsible for replication, recombination and repair (15.0%).

.

Function	Percentage of genes found in the pangenome	
	Core genes	Accessory genes
Translation, ribosomal structure and biogenesis	9.5	2.5
Transcription	6.3	9.4
Replication, recombination and repair	4.8	15.0
Chromatin structure and dynamics	0.1	1.2
Cell cycle control, cell division, chromosome partitioning	0.9	3.1
Defense mechanisms	1.2	2.2
Signal transduction mechanisms	3.2	8.5
Cell wall/membrane/envelope biogenesis	3.4	0.1
Cell motility	0.2	0.0
Intracellular trafficking, secretion, and vesicular transport	1.2	0.9
Posttranslational modification, protein turnover, chaperones	3.7	1.6
Energy production and conversion	7.4	2.8
Carbohydrate transport and metabolism	5.8	8.0
Amino acid transport and metabolism	9.5	6.6
Nucleotide transport and metabolism	4.6	1.7
Coenzyme transport and metabolism	6.2	2.4
Lipid transport and metabolism	2.7	2.3
Inorganic ion transport and metabolism	6.6	5.7
Secondary metabolites biosynthesis, transport and catabolism	1.5	1.6
General function prediction only	11.8	13.6
Function unknown	9.7	10.7

Table 5.12: COG family group of genes in the core and accessory genomes of *S. haemolyticus* isolates

The COG family group were identified for the unique genes found in the 10 *mecA*$^{+}$ *S. haemolyticus* isolates from East and West London that were not found in the other *S. haemolyticus* isolates from the ENA database (Table 5.13). 306 unique genes were identified of which the highest proportion of the genes have been identified to belong the COG group replication, recombination and repair (34.3%)

Function	**Percentage (%) of genes in the pangenome unique to *S. haemolyticus* isolates from East and West London**
Translation, ribosomal structure and biogenesis	0.2
Transcription	15.2
Replication, recombination and repair	34.3
Cell cycle control, cell division, chromosome partitioning	4.9
Defense mechanisms	2.4
Signal transduction mechanisms	0.6
Cell wall/membrane/envelope biogenesis	5.7
Intracellular trafficking, secretion, and vesicular transport	0.8
Posttranslational modification, protein turnover, chaperones	1.4
Carbohydrate transport and metabolism	9.3
Amino acid transport and metabolism	0.6
Inorganic ion transport and metabolism	3.5
General function prediction only	13.4
Function unknown	7.7

Table 5.13: Unique gene COG family from 10 *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings in East and West London.

Comparative analyses of the accessory genomes within *mecA*$^+$ *S. haemolyticus* isolates recovered from public settings in East and West London showed that East London isolates had 677 genes not found in those recovered from West London and that West London isolates had 676 genes not found in those recovered from East London. 7 out of these unique genes were ubiquitous in *mecA*$^+$ isolates recovered from East London, whereas 8 genes were ubiquitous in *mecA*$^+$ isolates recovered from West London. The known functions for the genes that were unique and found in all *mecA*$^+$ East London isolates were phosphomethylpyrimidine kinase, 3-hexulose-6-phosphate synthase, putative protein, long-chain fatty acid CoA ligase. The known function of unique genes found in *mecA*$^+$ West London isolates was tRNA-Val (tac), 3-hexulose-6-phosphate synthase, MFS family major facilitator transporter, proline/betaine: cation symporter and phosphomethylpyrimidine kinase. The difference in the accessory genome for the *mecA*$^+$ general public settings and *mecA*$^+$ public setting in hospital isolates was that general public setting isolates had 392 genes not found in isolates from public areas in hospitals, whereas from public areas in hospitals had 958 genes that were not found in public areas in the community isolates. For the isolates from the *mecA*$^+$ general public settings, there were 9 genes that were ubiquitous in these isolates where *mecA*$^+$ public area in hospitals isolates had 6 genes that were ubiquitous in these isolates but were not found in isolates from general public settings. The known function of ubiquitous genes in the general public settings was bifunctional acetyltransferase/phosphotransferase, *ISSau3* transposase, GNAT family acetyltransferase, transposase for *IS431mec*. The known functions of ubiquitous genes in the hospital isolates were recombinase/resolvase; Alpha/beta hydrolase fold-3 domain-containing protein, serine-type D-Ala-D-Ala carboxypeptidase and a replication initiator protein A. *MecA*$^+$ East London had a significantly higher proportion of unique genes compared with *mecA*$^+$ West London

isolates for replication, recombination and repair (22.8% and 14.0% respectfully (P= <0.0001) and amino acid transport and metabolism (7.7% and 4.1% respectfully (p= 0.0054)) (Table5.13). For West London isolates there was a significantly higher proportion of unique genes compared to *mecA*$^+$ East London isolates that had functions for carbohydrate transport and metabolism (10.4% and 4.6% respectfully (P=0.0001)); translation, ribosomal structure and biogenesis (3.3% and 1.0% respectfully (P=0.0047)); defense mechanisms (5.2% and 1.8% respectfully (P=0.0008)); carbohydrate transport and metabolism (10.5% and 4.6% respectfully (P=0.0001)); inorganic ion transport and metabolism (8.8% and 5.2% respectfully (P= 0.0106)) and secondary metabolism biosynthesis, transport and catabolism (2.6% and 0.9% respectfully (P=0.0183)). For *mecA*$^+$ general public setting isolates there was a significantly higher portion of unique genes that were responsible for replication, recombination and repair compared with *mecA*$^+$ hospital isolates from public area (33.2% and 13.4% respectfully (P=0.001)); cell cycle control, cell division, chromosome partition (2.5% and 1.4% respectfully (P=0.0423) compared with *mecA*$^+$ hospital isolates from public areas (Table 5.14). *MecA*$^+$ hospital isolates from public areas had a significant higher proportion of unique genes compared to *mecA*$^+$ general public settings isolates that had functions for carbohydrate transport and metabolism (9.7% and 3.8% respectively (P=0.0003)); defense mechanisms (4.5% and 1.1% respectfully (P=0.0023)); amino acid transport and metabolism (5.8% and 2.3% respectively (P=0.0052)); coenzyme transport and metabolism (3.1% and 0% respectively (P=0.0005)); Intracellular trafficking, secretion, and vesicular transport (1.3% and 0% respectively (P=0.0254)) and secondary metabolites biosynthesis, transport and catabolism (2.8 and 0% respectively (P=0.0009) (Table 5.15)

	Percentage (%) of isolates genes in the pangenome that were unique to isolates recovered from public settings	
Function	**East London**	**West London**
Translation, ribosomal structure and biogenesis	1.0	3.2
RNA processing and modification	0.0	0.0
Transcription	9.9	9.4
Replication, recombination and repair	22.8*	14.0
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	1.7	1.1
Nuclear structure	0.0	0.0
Defense mechanisms	1.8	5.2*
Signal transduction mechanisms	1.4	0.0
Cell wall/membrane/envelope biogenesis	7.0	8.6
Cell motility	0.0	0.1
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	0.5	1.1
Posttranslational modification, protein turnover, chaperones	0.8	0.5
Energy production and conversion	1.9	1.8
Carbohydrate transport and metabolism	4.6	10.4*
Amino acid transport and metabolism	7.7*	4.1
Nucleotide transport and metabolism	2.4	1.9
Coenzyme transport and metabolism	2.8	3.1
Lipid transport and metabolism	2.2	2.2
Inorganic ion transport and metabolism	5.2	8.8*
Secondary metabolites biosynthesis, transport and catabolism	0.9	2.6
General function prediction only	12.8	12.9
Function unknown	12.5	7.9

Table 5.14: Difference in COG function of the unique genes in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from East and West London*= The area where there was significantly higher ($p=<0.05$) percentage of genes for that COG function

	Percentage of isolates genes in the pangenome that were unique to Isolates recovered from public settings	
Function	**Community**	**Hospital**
Translation, ribosomal structure and biogenesis	1.9	2.9
RNA processing and modification	0.0	0.0
Transcription	9.2	9.9
Replication, recombination and repair	33.2*	13.4
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	2.5*	1.0
Nuclear structure	0.0	0.0
Defence mechanisms	1.1	4.5*
Signal transduction mechanisms	2.6	1.4
Cell wall/membrane/envelope biogenesis;	8.7	7.7
Cell motility	0.0	0.1
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	0.0	1.3
Posttranslational modification, protein turnover, chaperones	0.8	0.3
Energy production and conversion	1.1	2.1
Carbohydrate transport and metabolism	3.8	9.7*
Amino acid transport and metabolism	2.3	5.8*
Nucleotide transport and metabolism	2.9	1.9
Coenzyme transport and metabolism	0.0	3.1*
Lipid transport and metabolism	1.0	2.6
Inorganic ion transport and metabolism	5.1	7.8
Secondary metabolites biosynthesis, transport and catabolism	0.0	2.8
General function prediction only	13.0	13.0
Function unknown	10.7	8.6

Table 5.15: Difference in COG function of the unique genes in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public setting and West London.*= The area where there was significantly higher (p=<0.05) percentage of genes for that COG function

5.4 Discussion

5.4.1 *S. haemolyticus:* Virulence genes and HGT

S. haemolyticus is the second most common CoNS nosocomial infection behind *S. epidermidis* (Czekaj et al., 2015). The virulence genes can be acquired horizontally, but little is known of the origins of these genes in *S. haemolyticus* (O'Riordan and Lee, 2004). One key virulence gene in *S. haemolyticus* is the polysaccharide capsule genes which were identified in all *mecA*$^{+}$ isolates recovered from public settings and all isolates from ENA database used in the pangenomic analyses carried out in this study. Polysaccharide capsules have been previously reported in *S. haemolyticus* as being important for bacterial survival during infection as they impede phagocytosis (O'Riordan and Lee, 2004). This function in *S. haemolyticus* JCSC 1435 has been demonstrated to be encoded by an operon comprised of 13 genes in tandem. However, the majority of the isolates recovered from East and West London had only two capsules genes; only two isolates carried 14 polysaccharide capsule genes (Flahaut et al., 2008). The carriage of these genes was compared to the isolates obtained from the ENA database; the highest number of capsule genes identified was 18, which were mostly found in clinical isolates (isolates 115601, 51-30, 83131A, 83131B, ERR085179, ERR085182, ERR085183, SRR1182428, SRR1182429, SRR1182430, SRR1182431, SRR1182432) . This suggests that these additional capsule genes may make *S. haemolyticus* more virulent, whereas it was predict isolates with only 2 capsule gene will lack the ability to produce a capsule as each gene has been shown to code for different functions in producing polysaccharide capsules (Flahaut et al., 2008). These functions include chain length determination; putative tyrosine-protein kinase, putative phosphotyrosine-protein phosphatase, putative 4,6-dehydratase, gylocolesis transferase and amino transferase. Interestingly, from this analyses of *S. haemolyticus* JCSC 1435 there was 14 capsules genes, not 13 capsule genes

as reported by Flahaut and co-workers. (Flahaut et al., 2008) This extra capsule gene would require further laboratory work to understand if it was involved in capsule production along with isolates that are demonstrated to have additional polysaccharide capsule genes. Two isolates (445 and 538) recovered from West London Hospital in this study possessed 14 capsule genes of which 9 were predicted to be horizontally transferred. These genes were predicted to be donated from *Nosocomiicoccus*, *Listeria*, *Bacillus* and *Caloramator* genus. Previous reports have described a polysaccharide capsule in *Bacillus* genus but not in *Listeria*, *Nosocomiicoccus* or *Caloramator* genus (Scarff et al., 2018; Vázquez-Boland et al., 2001) It has been noted that *Bacillus cereus* and *Bacillus anthracis* has a multi-gene operon within a plasmid which encodes for polysaccharide capsules (Scarff et al., 2018) This would suggest the possible route of transmission of these genes in *S. haemolyticus* could be via plasmid conjugation. As mentioned in chapter four, these predicted donors are just the best match distal organisms and are most likely that the actual donor was an ancestor or an extinct donor (Zhu et al., 2014).

Other capsule genes that were ubiquitous to the *S. haemolyticus* isolates recovered from East and West London were *capB* and *capC* which encode the polyglutamic acid capsule and are responsible for immune evasion. These genes were identified in 100% of the isolates recovered from the public settings and in 86.8% of the isolates used in the pangenomic analysis. Similar to the *S. epidermidis* analyses, the *capA* and *capD* genes, were absent in *S. haemolyticus*. Although *CapB* and *capC* genes have not previously been identified as a virulence factor in *S. haemolytic*us, they have been identified as such in *S. hominis* (Calkins et al., 2016). Interestingly, a report has shown that *S. haemolyticus* surface-associated protein reacts with antibodies of polyglutamic acid, but no studies have shown that they are involved in bacterial virulence (Flahaut et al., 2008). These

antibodies may be able to bind to other proteins in *S. haemolyticus*, which may have a similar structure to that of polyglutamic acid.

The *nuc* gene which encodes the thermonuclease was identified in all *S. haemolyticus* isolates recovered from public settings and the isolates obtained from the ENA database. This suggests that this gene is essential for *S. haemolyticus* survival than for its virulence. This gene has been characterised in *S. aureus* to hydrolyse DNA and RNA and is involved in host cells evasion of neutrophil extracellular traps as well as inhibition of biofilm formation via cleavage of extracellular DNA (Kiedrowski et al., 2014).

In this study, genes were identified that have not previously been identified in *S. haemolyticus*, including the *cylR2* gene, which encoded the cytolysin regulator in *Enterococcus* spp. *CylR2* was found in 100% of *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public setting in East and West London. In addition, it was also found in 95.2% of S. *haemolyticus* isolates used in the pangenome analysis. These genes may have been missed as the predicted protein structure has low sequence similarity to that of known cytolysin, but they are shown to have close paralog as described by the VFanalyzer software (Liu et al., 2019). Additionally, this gene is shown to be part of an 8 gene operon in *Enterococcus* spp. Therefore, the absence of the other 7 genes in *S. haemolyticus* would suggest that they would lack the ability to produce cytolysin (Van Tyne et al., 2013). From HGT analyses of the *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings from East and West London, the *cylR2* gene was predicted to be horizontally transferred with the closet donor link identified as *Streptococcus* and *Virgibacillus* genus that have previously been reported to produce cytolysin (Molloy et al., 2015).

From this study, genes found in the in *icaADBC* operon that are responsible for biofilm production in *S. epidermidis* were not found in the *S. haemolyticus* isolates. Fredheim and

co-workers found that 53 out of 72 (74%) clinical *S. haemolyticus* isolates produce biofilm, though only 2 isolates had the *icaADBC* operon (Fredheim et al., 2009). This shows that *icaADBC* operon is not essential in biofilm production in *S. haemolyticus* as in the case of *S. epidermidis*; therefore, the bacterium uses a different mechanism to produce biofilm, which is yet unknown. These findings in this report would suggest that it is possible that the *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public setting could also produce biofilm although further laboratory experiments will be required to confirm this hypothesis.

No clustering was observed based on virulence gene profiles or on their isolation sources, however, virulence gene profiles of the isolates recovered from public settings were similar to those obtained from the ENA database that has previously been recovered by others from clinical samples, health humans, livestock and companion animals and further ENA database isolates that were recovered from public settings and natural environments. Although virulence phenotype was not determined in this study, these isolates have potential to cause infection in humans and are a public health risk as they are similar virulent genes to that found in known clinical isolates that have cause blood, eye, sputum and colon infections.

5.4.2 *S. haemolyticus:* Antibiotic resistance genes and HGT

Understanding which antibiotic resistance genes are horizontally transferred; their mode of transfer and their origins is important to understand bacterial evolution. In this study, all the genes, except for the beta-lactamase encoding gene *blaZ* were predicted to be horizontally transferred in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings. The *blaZ* gene was carried on the plasmids of 6 isolates. The method used to detect gene transfer can only predict if the gene originated from a different genus (Zhu et al., 2014).

Therefore, these genes may be transferred from another *S. haemolyticus* isolate or other staphylococcal species. The predicted donor link for the antibiotic resistance genes was the same as the majority of antibiotic resistance genes in *S. epidermidis* analyses in chapter 4. This would suggest that antibiotic resistance genes that were horizontally transferred have the same ancestral link in CoNS *mecA*$^{+}$ isolates from public settings in East and West London. The only antibiotic resistance genes in *S. haemolyticus* isolates in which the donor species were different were *AAC(6')-Ie-APH(2'')-Ia* and *APH(3')-IIIa*. The predicted donor for *AAC(6')-Ie-APH(2'')-Ia* gene for isolates 1 and 105 was *Clostridiales bacterium VE202-16* and for isolate 93 *Streptococcus mitis B6* and isolate 99 was and *Enterococcus faecalis V583*. The predicted donor for *APH(3')-IIIa* for isolate 93 was *Streptococcus mitis B6* and isolate 99 was *Enterococcus faecium DO. Streptococcus mitis, Enterococcus faecalis and Enterococcus faecium* are three known pathogens that have an array of infection from endocarditis, urinary tract infection and even meningitis. This demonstrates that *S. haemolyticus* could horizontally acquire antibiotic resistance genes from other pathogens from the public environment. The *AAC(6')-Ie-APH(2'')-Ia* and *APH(3')-IIIa* genes were only found in the *mecA*$^{+}$ isolates in East London public settings which suggests that there is a pool of bacterial species within this area that has the potential of transferring the same antibiotic resistance genes into *S. haemolyticus*.

Among the *mecA*$^{+}$ *S. haemolyticus* isolates from public settings, there were 3 isolates (1, 105 and 405) that had the majority of their antibiotic resistance genes carried on their plasmids with one isolate (405) shown to have at least 5 genes encoded on a single plasmid. Bacteria plasmids are very common way of exchanging genetic information to other bacteria. By picking multiple genes on a plasmid and then picking multiple plasmids, the bacterium becomes multidrug resistant (Millan, 2018).

Five antibiotic resistance genes were identified within genomic island regions in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings. These included *blaZ*, *APH(3')-IIIa*, *fusB*, *msrA*, *mphC*, *ermA, AAC(6')-Ie-APH(2'')-Ia* and *tet(K)*. As with *mecA*$^{+}$ *S. epidermidis* isolates from public settings, the *mecA* genes was not found within the predicted genomic island regions in *mecA*$^{+}$ *S. haemolyticus* isolates from public settings. There is only one previously published report characterising a genomic island in an *S. haemolyticus* isolate (JCSC1435). The authors of this study found that the *mecA*, blaZ, *msrA*, *mphC, AAC(6')-Ie-APH(2'')-Ia* antibiotic resistance genes were found within known genomic islands. These islands were associated with the SCC, transposons and integrated plasmids (Takeuchi et al., 2005).

Unlike *mecA*$^{+}$ *S. epidermidis* isolates from public areas in East and West London, antibiotic resistance genes were identified within the phage insertion sequence in 2 out of the 10 *S. haemolyticus mecA* isolates. These genes were *APH(3')-IIIa* and *fusB*. Previous studies have shown that *the fusB* gene was found in a phage-related island in *S. epidermidis* (Chen et al., 2013). Little is known about CoNS of antibiotic resistance genes transfer by phages. This may be because phage transduction of antibiotic resistance genes is rare (Enault et al., 2017) Nevertheless, it would be interesting to know if a particular antibiotic resistance gene was transferred via a phage more often than others. For completeness of the phage insertion in the host genome, there were only three isolates (93, 99 and 361) out of the 10 *mecA*$^{+}$ *S. haemolyticus* isolates from public setting had intact phage insertion sequences. However, other *S. haemolyticus mecA*$^{+}$ isolates recovered from public settings to have incomplete phage sequences. This would suggest that many of these phages that were found 'incomplete' had their genes deleted by the host genome as they do not confer any advantage for bacterial survival and may act as a metabolic burden (Ramisetty and Sudhakari, 2019). Additionally, two isolates (445 and

492) had an incomplete phage insertion sequence was associated with *Synechococcus* and *Prochlorococcus* species. As mentioned in the previous chapter, these sequences are most likely unknown phages that have not previously been studied. The intact phage sequence has not been previously identified in *S. haemolyticus* but has been found in *S. hominis* and *S. capitis* (*Staphylococcus phage StB12, Staphylococcus phage StB27*), and *S. epidermidis* (*Staphylococcus phage CNPx*) (Deghorain et al., 2012; Maniv et al., 2016). This shows that *Staphylococcus* phages can infect and integrate into different *Staphylococcus* species; hence, it is possible for virulence genes horizontally transferred via phages into *S. aureus* to also be transferred into *S. haemolyticus* isolates.

Many studies have described antibiotic resistance phenotypes in *S. haemolyticus,* but few studies focused on their antibiotic resistance genes profile (Cavanagh et al., 2014; Maleki et al., 2018). For the overall antibiotic resistance gene in all 83 *S. haemolyticus* isolates *mecA* (78.3%) and *blaZ* (72.3%) genes were the most common. This was similar to *S. epidermidis* except for the *norA* which is ubiquitous among *S. epidermidis* isolates but not present in *S. haemolyticus* isolates. The only antibiotic resistance gene found in this study that was not found in *S. epidermidis* isolates was *erm(33)*. This gene has previously only been identified in *S. aureus* and *S. sciuri* isolates (Li et al., 2015; Schwarz et al., 2002). Interestingly, *fosB3* (which encodes for resistance to fosfomycin) and *dfrC* (encoding resistance for diaminopyrimidine) were present only in the isolates from ENA database recovered from clinical samples suggesting their frequent use in clinics but not in other environments (Ltd, 2016).

A hierarchy clustering heatmap revealed no clustering for the antibiotic resistance gene profiles based on their isolation source. However, isolates recovered from public settings in East and West London had similar antibiotic resistance gene profiles when compared

to profiles of isolates obtained from the ENA database, including isolates recovered from clinical sources, healthy humans, livestock and companion animals and those recovered from other public settings and natural environments. As antibiotic resistance is higher in abundance in hospital environments and livestock animals due to high usage of antibiotics, it is expect that many of these isolates have crossed over into different niches as well as horizontally transferred their resistance genes to isolates found in public settings (Cantón and Morosini, 2011).

5.4.3 HGT in *mecA*$^+$ *S. haemolyticus* isolates recovered from East and West London

Horizontal transfer of antibiotic resistance and virulence genes is an important evolutionary adaption, but little is known about what proportion of the *S. haemolyticus* genome that has been donated from other organisms. In this study, it was predicted that 8.9% to 13.5% of the genes were HGT-derived from *mecA*$^+$ *S. haemolyticus* isolates recovered from East and West London. This was similar to that predicted for *S. epidermidis* isolates in this study (8.6% to 10.1%). There were also similarities in the predicted common genera which donated the majority of genes (*Bacillus* and *Macrococcus*). This would suggest that the S. *haemolyticus* accessory genome had similar evolution as in *S. epidermidis* isolates as the predicted donor species were commonly found in these environments (Conlan et al., 2012).

5.3.4 Phylogenetic and pangenome analysis of *S. haemolyticus* isolates recovered from public settings with *S. haemolyticus* isolates obtained from the ENA database.

S. haemolyticus can be found in different ecological niches but little is known about their genetic relatedness with each other (Barros et al., 2012; Ruzauskas et al., 2014; Seng et al., 2017b; Xu et al., 2015). Here, a phylogenetic analysis was carried out to compare the *mecA*$^+$ *S. haemolyticus* isolates recovered from public settings from East and West

London with isolates obtained from the ENA database that have previously been recovered from clinical samples (blood, colon, sputum and eye); healthy human (skin, nares and eyes); livestock (cows) and companion animal (dog); other isolates recovered from public settings and natural environment (groundwater), catheters from hospital and plant-associated isolates (Kefir seeds and willow). Similar to *S. epidermidis* phylogenetic analysis, isolates were split into two distinct clades, although those that were obtained from the ENA database and have been designated as 'clinical' were found in both clades. All but two isolates recovered from public settings in this study were identified as being genetically related to isolates recovered from clinical samples from an eye (SH1572), blood **(**M-176) and central venous catheter (95671) obtained from the ENA database. These results show that these isolates from public settings are a public health risk as they come from the same genetic lineages of isolate's that have shown to cause eye (conjunctivitis) and blood infections (Cavanagh et al., 2014; Panda and Singh, 2016). This includes the 4 *mecA*$^{+}$ isolates that were recovered from general public settings in East London (1, 93, 99 and 105). This would suggest that isolates that can cause blood infections are not just found in hospital areas but can be found in general public areas. Additionally, one isolate from West London hospital (492) was genetically related to isolates recovered from a dog (SW007) and another isolate from West London hospital (445) was phylogenetically related to an isolate recovered from kefir seeds (OG2) (Bean et al., 2017). Previous reports have indicated that companion animals are a potential reservoir for the *mecA*$^{+}$ *S. haemolyticus*, which could be transmitted to human via contact. Reports of *S. haemolyticus* recovered from plants is scarce but from this study analysis, isolate recovered from Kefir seeds harboured the *mecA* gene (Gan et al., 2014). In this study, it was not determined whether the *S. haemolyticus* isolates recovered from a dog could be a public health risk as there are no studies that have shown there to be a link

between genetic lineages of isolates recovered from companion animals and those that have caused human infections. However, there was a genetic relatedness between two isolates from ENA database recovered from cows' milk (NW19 and BC05211) and isolates recovered from clinical samples including isolate recovered from eye infection (SH1574) and blood infection (8074328), potentially posing public health risk. It has not been reported that *S. haemolyticus* from livestock belong to the same genetic lineages known to cause infections. Additionally, isolate from the ENA database recovered from plants (OG2), were also related to the isolate from the ENA database recovered from clinical blood samples (M-176). This shows that infective *S. haemolyticus* can be transmitted to humans via contact from plants or animals either as a foodstuff or direct contact with the actual plant or animal (Prado et al., 2015).

In this study, *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings in the community and public areas in hospitals from East London were genetically similar based on their accessory genome, and clustering, despite that they belonged to different core genome clusters (Figure 5.4). This indicates that these isolates have a shared pool of genes that may horizontally transfer due to similarities in bacterial species/strains found in the same geographical area (Segerman, 2012). The accessory genomes of seven *mecA*$^{+}$ *S. haemolyticus* isolates (1, 93, 99, 105, 506, 445 and 538) recovered from public settings were similar compared with accessory genomes of clinical isolates obtained from ENA database. These seven isolates also belonged to the same core genetic lineages as isolates recovered from clinical samples. This provides evidence that *S. haemolyticus* isolates recovered from public settings in this study are not simply harmless bacteria that picked resistance genes over the time, but most likely originated from the same niches and similarly possess genes typical for human pathogens capable of causing infections. This is also indicated by the virulence gene profiles as many of the isolates recovered public

settings in this study shared the same known virulence genes found in clinical isolates from the ENA database. Moreover, three isolates recovered from public settings (506, 492 and 361) were genetically different with the clinical isolates based on their core genome; however, they were similar based on their accessory genome to clinical isolates from ENA database recovered from vagina (DNF00585) sputum (C10F) and colon (1HT3). This would indicate that these three isolates originate from a different area to these clinical isolates but may have been transmitted to areas where clinical isolates were present in which they have horizontal acquired similar genes to each other in their accessory genome. Additionally, based on their accessory genomes, these isolates recovered from public settings were genetically similar to clinical isolates were also shown to be similar to isolates from ENA database recovered from plants, livestock and companion animals. This would indicate at one point these isolates recovered from livestock, companion animal and plants were transmitted to clinical areas, or clinical isolates has been transmitted outside of hospitals via humans. Therefore, these isolates may have horizontally acquired virulence genes from clinical isolates, or they are clinical isolates that have survived outside of hospital environments. As virulence determinants in staphylococci are found in the accessory genome, this suggests that these isolates recovered from plants, different environments, livestock and companion animals have similarities in their accessory genome with the clinical isolates, may potentially cause infection as they carry virulent genes that were responsible for initiating infection (Bosi et al., 2016). Additionally, there were 4 clusters that only had *S. haemolyticus* isolates from the ENA database that was recovered from clinical isolates (Figure5. 4). These isolates are most likely nosocomial and have not spread into different niches (Degener et al., 1994).

In a similar manner to the *S. epidermidis* pangenome analyses, it was observed that the *mecA*$^{+}$ *S. haemolyticus* from public settings and the *S. haemolyticus* isolates from the ENA database is an open pangenome in which it seems there is an unlimited number of genes as part of the species gene pool (Figure 5.4). This suggests that *S. haemolyticus* can be found in different environments within mixed microbial communities in which they can exchange genetic material to increase their gene pool (Bosi et al., 2015). Additionally, from the pangenome analysis of *S. haemolyticus* isolates, it was observed that there were no particular genes associated with different isolation sources (Figure 5.5). Despite this, it was observed a large proportion of the unique genes found in *mecA*$^{+}$ *S. haemolyticus* isolates from public areas from East and West London belonged to the COG family group that encode for the function of replication, recombination and repair (34.3%). This was more than what was identified in the *mecA*$^{+}$ *S. epidermidis* isolates recovered from East and West London (24.4%). This would suggest there were more novel genes acquired encoding these functions in *S. haemolyticus* then in *S. epidermidis*. As discussed in case of *S. epidermidis* recovered from public settings in this study, these unique genes were important for acquiring mobile genetic elements as they might be beneficial for the bacterial survival on not very 'desirable' high-frequency touched surfaces as these are not their preferred niche to live in/on. Interestingly, in this study there was a higher percentage of unique genes encoding for replication, recombination and repair in *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings in the community (33.1%) compared to that of isolates recovered from public areas in hospitals (13.4%). This indicates that the *mecA*$^{+}$ *S. haemolyticus* isolates from public settings may experience more DNA damage than to those found in public areas in hospitals. In public settings the bacteria may be exposed more to UV light, ionizing radiation and genotoxic chemicals that can cause DNA damage compared to bacteria found in hospitals (Žgur-Bertok, 2013).

Additionally, hospital isolates had significantly more unique genes responsible for functions such as carbohydrate transport and metabolism; defense mechanisms, amino acid transport and metabolism; coenzyme transport and metabolism, secondary metabolites biosynthesis, transport and catabolism. Five out of six of these functions is for transport or metabolisms which suggests that within hospital areas there is a larger amount of carbohydrate, amino acid and molecules for coenzymes than in public areas in the hospital. The higher number of unique genes that are for defense mechanisms might suggest there is a higher abundance of phages in hospitals than in public settings. Some genes that were ubiquitous to *mecA*$^{+}$ *S. haemolyticus* isolates from public settings that were absent in the *mecA*$^{+}$ isolates in hospitals and *vice versa*. The genes that were found to be unique in the *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings were most likely not essential for survival in these niches but may be different because of horizontal transfer of genes from different bacterial species found in these different areas. As in this study, there was only 10 *mecA*$^{+}$ *S. haemolyticus* isolate that was WGS, therefore if there were additional sequenced *S. haemolyticus* isolates from these areas, it would be expected not find genes that were ubiquitous and unique to isolates from general public setting compared to isolates recovered from public areas in hospitals and *vice versa*.

5.5 Conclusion

In this chapter, public setting *mecA*$^{+}$ *S. haemolyticus* isolates from East and West London were genetically compared to *S. haemolyticus* isolates recovered from clinical samples (eye, blood, sputum, colon), healthy humans skin and nares, eye; livestock (cows) and companion animal (dog), hospital environment from central venous catheters, various public settings and the natural environment, including from plants form the ENA database. From the phylogenetic analysis, the genetic relatedness of *S haemolyticus* isolates from public settings was determined to that of other isolates from the ENA database which has

not previously done before. Additionally, the antibiotic resistance genes, mobile genetic elements, virulent genes and pangenome were compared. From these analyses, there were novel findings. These findings were:

1. *S. haemolyticus* isolates have a variable number of genes that may be involved in polysaccharide capsule production.

2. *S. haemolyticus mecA*$^{+}$ isolates recovered from public settings were shown to be the same genetic lineages of isolates from the ENA database recovered from clinical samples from an infected eye and blood as well as being similar in their virulence gene profiles. This shows that *mecA*$^{+}$ *S. haemolyticus* isolates from public settings pose a potential public health risk.

3. *MecA*$^{+}$ *S. haemolyticus* isolates recovered public areas in the community and public areas in hospitals in East London were similar in their accessory genomes suggesting they horizontally acquired similar genes due to the similarities in the microbiome in that geographical area.

4. *MecA*$^{+}$ *S. haemolyticus* isolates recovered from general public settings possessed more unique genes encoding for replication, recombination and repair compared to *mecA*$^{+}$ *S. haemolyticus* isolates recovered from hospitals. This suggests that there are environmental factors such as UV light, ionizing radiation and genotoxic chemicals that could cause DNA damage to the bacteria found in general public settings than in public areas in hospitals.

Chapter 6: Comparative genomics of *mecA* positive *S. hominis* isolates recovered from public settings using the One Health approach; determining horizontal gene transfer in these isolates

6.1 Introduction

Staphylococcus hominis is the third most common coagulase-negative *Staphylococcus* (CoNS) infection from clinical cases which has been shown to cause bacteraemia, septicaemia, endophthalmitis, and endocarditis (Chaves et al., 2005; Cunha et al., 2007; Iyer et al., 2005). Pathogenicity studies in *S. hominis* are limited, but they have been characterised to produce biofilms, adhere to host cell, immune invasion and the activity of extracellular toxins (Szczuka et al., 2018). Previous reports have also identified them to be resistant to multiple antibiotics, but there are no previous studies that have used whole genome sequencing (WGS) to understand phylogenetic relatedness of *S. hominis* from public settings to isolates that have been previously recovered from clinical samples, healthy humans, livestock and other animals, natural environment and plants (Szczuka et al., 2018; Xu et al., 2018b).

6.2 Method

In this chapter, comparative genomic analyses were performed on the *mecA* positive (*mecA*$^{+}$) *S. hominis* isolates recovered from public settings in East and West London. This included identification of horizontal gene transfer (HGT) within the genome and the potential donor organisms of horizontally transferred antibiotic resistance and virulence genes. Additionally, it was investigated whether virulence or antibiotic resistance genes were transferred via plasmids, phages or by other mobile genetic elements. A phylogenetic tree was constructed to determine the genetic relationship of *mecA*$^{+}$ isolates from this study with *S. hominis* isolates using reference and draft whole genome sequenced (WGS) isolates from the European Nucleotide Archive (ENA) database that was previously recovered from different sources, including clinical samples (blood),

healthy human skin, livestock (cows), mosquitos, from natural environment and plants. Additionally, a pangenome was constructed to identify the core and accessory genes in the genome. Additionally, the Clusters of Orthologous Groups (COG) function family was compared between the unique genes found in *mecA*$^{+}$ *S. hominis* isolates collected from public settings in East and West London as well as general public settings and public areas in hospitals.

6.3 Results

6.3.1 Identification of Virulence genes in isolates recovered from public settings in East and West London

Twelve virulence genes were predicted in the 10 *mecA*$^{+}$ *S. hominis* isolates from East and West London based on the VFanalyzer software (Table 6.1). The virulence genes that were identified in all the isolates were the *atl* gene (encoding autolysin involved in bacterial adhesion); *lip* gene (encoding lipase enzyme which is involved in the detachment of bacterial cells from colonised sites); *nuc* gene (encoding thurmonuclease an enzyme that can hydrolyse the host cell DNA and RNA), capsule genes (involved in immune invasion) *capB* and *capC* gene (encoding for polyglutamic acid capsule for immune invasion). *WbtE* gene (encoding for lipopolysaccharide modification in *Francisella* spp. necessary for immune invasion) was identified in 20% of the isolates. *icaA*, *icaB*, *icaC* genes (encoding for intercellular adhesion proteins involved in biofilm formation) and *cylR2* gene (encoding cytolysin which is involved in lysing erythrocytes, polymorphonuclear leukocytes and macrophages) were found in 10% of the isolates. Interestingly, isolate 207 and 209 had 6 copies of the capsule genes, whereas isolate 479 had 15 copies. 207 and 209 were predicted to have two copies of polysaccharide capsule genes. A hierarchical cluster heatmap of virulence gene profiles of public setting isolates showed no clustering of *mecA*$^{+}$ *S. hominis* isolates based on whether they were recovered

from West London or East London or whether they were recovered from general public settings or public areas in hospitals (Figure 6.1).

Virulence genes	**Function**	**Percentage (%) of the *mecA* isolates from public areas East and West London**
nuc	Thermonuclease	100
lip	Lipase enzyme	100
capsule	Capsule genes	20.0
atl	Autolysin	100
capB	Polyglutamic acid capsule	100
capC	Polyglutamic acid capsule	100
cylR2	Cytolysin enzyme	10
wbtE	lipopolysaccharide modification	20.0
icaA	intercellular adhesion proteins	10.0
icaB	intercellular adhesion proteins	10.0
icaC	intercellular adhesion proteins	10.0
Polysaccharide capsule	Polysaccharide capsule	20.0

Table 6.1: Percentage of Virulent genes found in the 10 *mecA*$^{+}$ *S. haemolyticus* from East and West London

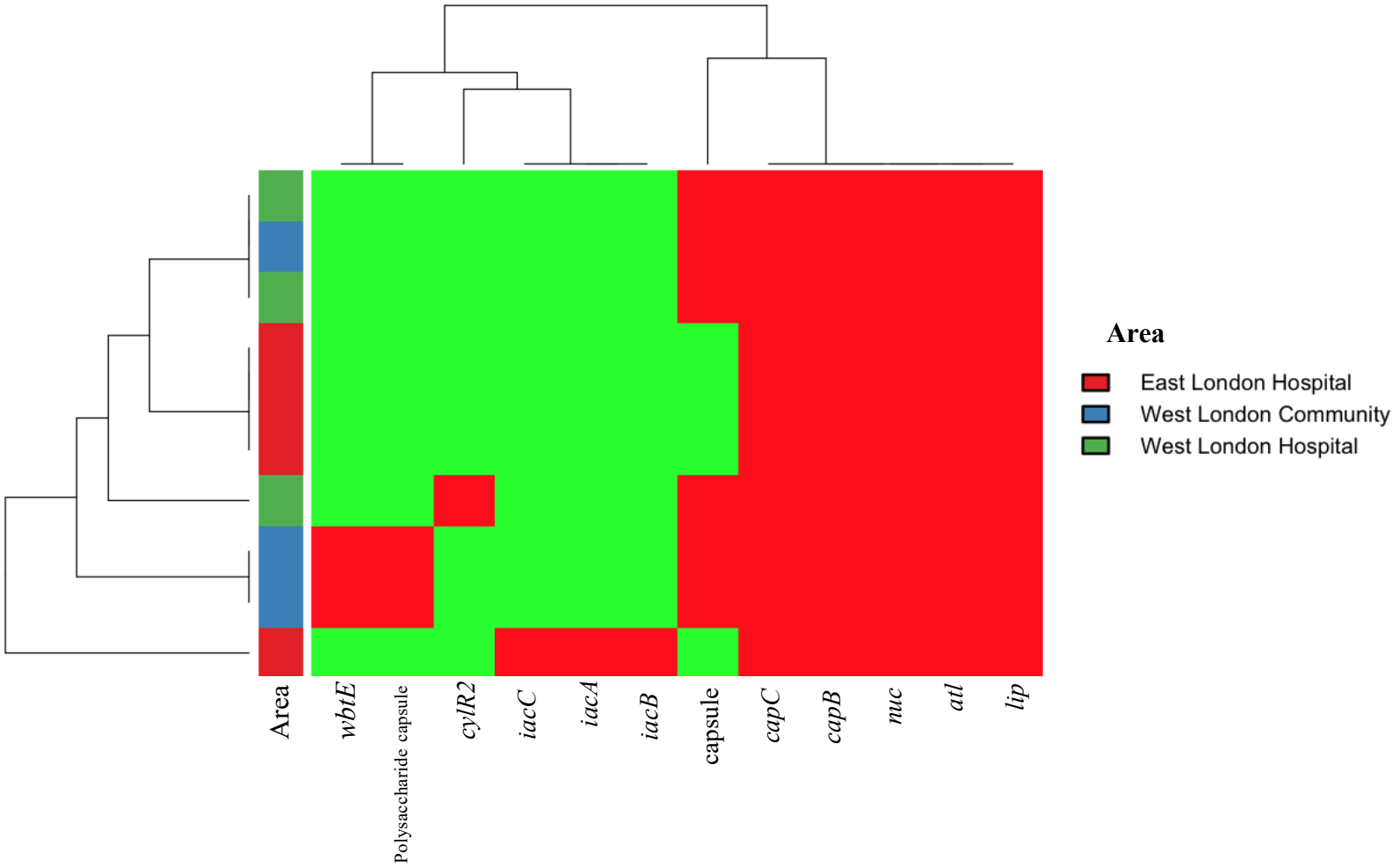


Figure 6.1: Hierarchy cluster heatmap of virulence gene profiles of *mecA*$^{+}$ S. *hominis* isolates from public areas in East and West London. Red tile gene present; Green tile gene absent

6.3.2. Horizontal gene transfer events in *mecA*$^+$ *S. hominis* isolates recovered from public settings in East and West London

Horizontal gene transfer was predicted in the 10 *mecA*$^+$ *S. hominis* isolates from public settings in East and West London and in the reference *S. hominis* isolate K1 using the HGTector pipeline (Table 6.2). Genes that were horizontally transferred were predicted in 7.8% to 9.5% of the *mecA*$^+$ isolates from public settings in East and West London compared to the reference isolate K1 (recovered from a cow) in which 9.4% of its genome was predicted to be horizontally transferred based on BLAST hit distribution patterns from the NCBI non-redundant protein sequences database (Pruitt et al., 2007; Zhu et al., 2014). These genes were predicted to be donated from 99 different genera based on the best hit of the non-redundant protein sequences database and the NCBI taxonomy database. The genera predicted to have most of the genes were donated from *Salinicoccus* (mean n=27); *Bacillus* (mean n=26) and *Macrococcus* (mean n=23) (Table 6.3).

S. hominis ID	Number of chromosomal protein-coding genes	Number of predicted HGT-derived genes	Percentage of HGT derived genes (%)
K1	2192	206	9.4
207	2180	194	8.9
208	2163	191	8.8
209	2157	188	8.7
372	2084	181	8.7
385	2108	190	9.0
386	2132	194	9.1
387	2161	189	8.8
479	2140	204	9.5
620	2033	159	7.8
623	2085	181	8.7

Table 6.2:**Number of HGT genes in *mecA*$^+$ *S. hominis* from East and West London and in the reference isolate *S. hominis* K1**

	S. hominis ID										
Predicted donor genus	**385**	**372**	**387**	**620**	**207**	**623**	**479**	**208**	**209**	**386**	**K1**
Acidibacillus	0	0	0	0	0	0	0	0	0	1	0
Aerococcus	1	1	1	1	2	1	1	1	2	1	2
Alysiella	0	0	0	0	0	0	0	0	0	0	1
Amphibacillus	0	0	0	0	1	0	0	0	1	0	1
Anaerobacillus	0	1	2	1	2	1	2	2	2	2	1
Anoxybacillus	0	1	1	1	1	1	0	0	1	1	0
Anthococcus	0	1	1	1	1	1	1	1	1	1	1
Atopococcus	1	1	1	1	1	1	1	1	1	1	1
Atopostipes	0	0	0	0	0	0	1	0	0	1	0
Auricoccucs	4	5	5	4	4	5	5	3	3	4	6
Bacillus	26	25	26	23	22	27	26	28	23	29	26
Bariatricus	1	0	0	0	0	0	0	0	0	0	0
Blautia	1	0	0	0	0	0	0	0	0	1	0
Brevibacillus	5	5	5	5	4	5	4	2	4	5	5
Burkholderia	0	0	1	0	0	0	0	0	0	0	0
Carnobacterium	4	2	2	2	3	2	3	4	3	2	5
Caryophanon	0	0	0	0	0	0	0	0	0	0	1
Chloracidobacterium	0	0	1	0	0	0	0	0	0	0	0
Clostridium	1	1	1	0	1	1	2	2	1	0	1
Corynebacterium	0	0	1	0	0	0	0	0	0	1	0
Cutibacterium	3	3	5	3	4	3	3	3	4	4	2
Desmospora	1	0	0	0	0	0	0	0	0	0	0
Domibacillus	0	1	2	1	0	1	0	0	0	2	0
Dorea	0	0	0	0	0	0	1	0	0	0	0
Edaphobacillus	0	0	0	0	0	0	0	1	0	0	0
Enteractinococcus	0	0	1	0	0	0	0	0	0	1	0
Enterobacter	0	0	0	0	1	0	1	0	1	0	0
Enterococcus	9	9	2	6	10	9	9	5	9	4	4
Eremococcus	1	1	1	1	1	1	0	1	1	0	1
Exiguobacterium	2	1	1	1	1	1	2	1	1	1	1
Faecalibaculum	0	0	0	0	0	0	0	1	0	0	0
Faecalicatena	0	0	0	0	0	0	0	1	0	0	0
Fictibacillus	1	1	1	1	1	1	1	1	1	1	1
Fusobacterium	0	0	1	0	0	0	0	0	0	0	0
Gemella	0	0	0	0	0	0	0	0	0	1	0
Geobacillus	1	1	1	1	1	1	1	1	1	1	1
Gracilibacillus	0	6	3	6	1	6	0	1	1	5	0
Gulosibacter	0	0	0	0	0	0	0	0	0	0	1
Halarchaeum	1	1	1	1	1	1	1	1	1	1	1
Halobacillus	0	1	1	1	2	1	0	0	2	1	0
Halogranum	1	0	0	0	0	0	0	0	0	0	0
Halopiger	0	0	0	0	0	0	1	0	0	1	0
Hathewaya	0	0	1	0	0	0	0	0	0	0	0
Herbaspirillum	0	0	0	0	0	0	0	1	0	0	0
Jeotgalibacillus	0	0	0	0	0	0	0	1	0	0	1
Jeotgalicoccus	7	5	6	3	2	5	10	4	2	4	7

	S. hominis ID										
Predicted donor genus	**385**	**372**	**387**	**620**	**207**	**623**	**479**	**208**	**209**	**386**	**K1**
Kroppenstedtia	0	0	0	0	0	0	0	0	0	0	1
Kurthia	0	1	2	1	3	1	0	1	3	2	1
Lachnoanaerobaculum	1	0	0	0	0	0	0	1	0	0	0
Lactobacillus	9	8	9	6	9	8	9	7	9	9	10
Lentibacillus	0	0	0	0	0	0	4	7	0	0	2
Leptotrichia	1	0	0	0	0	0	0	0	0	0	0
Listeria	3	3	3	3	5	3	8	4	5	3	3
Lysinibacillus	3	5	5	5	5	5	3	2	5	5	2
Macrococcus	22	24	18	21	26	24	24	25	26	24	34
Marinilactibacillus	0	0	0	0	0	0	0	0	0	1	0
Marinococcus	1	1	0	1	0	1	1	1	0	0	1
Massilibacterium	1	1	1	1	1	1	1	1	1	1	1
Micrococcus	0	0	0	0	0	0	0	0	0	0	1
Mycoplasma	0	0	1	0	0	0	0	0	0	0	0
Neisseria	0	0	0	0	0	0	1	0	0	0	0
Nocardia	0	0	0	0	0	0	0	1	0	1	0
Nosocomiicoccus	3	1	1	1	4	1	6	2	4	2	8
Oceanobacillus	7	7	9	7	10	7	8	10	10	10	10
Oenococcus	0	0	0	0	1	0	1	0	1	0	0
Oribacterium	1	2	2	2	2	2	2	2	2	2	2
Ornithinibacillus	0	1	0	1	2	1	1	1	2	0	0
Paenibacillus	4	4	6	2	3	3	3	2	2	3	5
Paucisalibacillus	0	0	0	0	0	0	0	0	0	0	1
Pediococcus	0	0	0	0	0	0	0	0	0	0	1
Peptoclostridium	1	1	1	1	1	1	1	1	1	1	1
Phormidesmis	0	0	0	0	0	0	0	0	0	0	1
Planococcus	2	2	4	2	3	2	2	2	3	4	3
Pontibacillus	0	0	0	0	0	0	1	0	0	0	0
Providencia	0	0	0	0	0	0	0	1	0	0	0
Pseudomonas	5	0	0	0	0	0	0	0	0	0	4
Rummeliibacillus	0	0	0	0	0	0	0	0	0	0	1
Saccharibacillus	0	0	0	0	0	0	0	1	0	0	0
Salimicrobium	0	0	0	0	1	0	0	0	1	0	0
Salinicoccus	26	26	28	21	31	26	26	30	26	27	18
Salsuginibacillus	1	1	0	1	0	1	0	0	0	0	1
Solibacillus	1	0	0	0	0	0	0	0	0	0	1
Sporolactobacillus	2	1	0	1	0	1	1	0	0	1	0
Sporosarcina	3	4	2	4	3	3	4	2	3	4	3
Streptococcus	8	6	9	6	6	6	8	7	7	7	5
Streptomyces	0	0	0	0	1	0	4	0	1	0	2
Sulfobacillus	1	0	0	0	0	0	0	0	0	0	0
Terribacillus	1	1	1	1	1	1	1	1	1	1	1
Tetragenococcus	2	1	1	1	2	1	2	2	2	1	1
Thalassobacillus	0	0	2	0	1	0	0	1	1	1	1
Thermoactinomyces	0	0	0	0	0	0	0	0	0	0	1
Trichococcus	0	1	1	0	0	1	0	0	0	2	0
Tuberibacillus	0	0	1	0	1	0	0	0	1	0	0
Tumebacillus	0	0	0	0	0	0	0	1	0	0	0

	S. hominis ID										
Predicted donor genus	**385**	**372**	**387**	**620**	**207**	**623**	**479**	**208**	**209**	**386**	**K1**
Vagococcus	1	1	1	1	1	1	1	1	1	1	1
Veillonella	1	0	2	0	1	0	0	3	1	0	0
Virgibacillus	4	1	1	1	0	1	1	1	0	1	2
Viridibacillus	0	0	0	0	0	0	0	0	0	0	1
Vulcanibacillus	0	0	0	0	1	0	0	0	1	0	0

Table 6.3: *MecA*$^{+}$ *S. hominis* isolates from public settings in East and West London and reference *S. hominis* K1 HGT-derived genes by their putative donor genus as indicated by the best distal match.

Ten out of 12 antibiotic resistance genes were predicted to be horizontally transferred from other organisms (Table 6.4). The only genes not predicted to be horizontally transferred were *ANT(4')-Ib* and *APH(3')-IIIa*.

S. hominis ID	Antibiotic resistance genes	Closest predicted donor species
207	*blaZ*	*[Propinobacterium] namnetense SK182B-JCVI*
	mecA	*Macrococcus canis*
	qacA/B	*Bacillus ndiopicus*
	msrA	*Veillonella atypica*
	fusB	*Enterococcus thailandicus*
208	*blaZ*	*[Propinobacterium] namnetense SK182B-JCVI*
	mecA	*Macrococcus canis*
	msrA	*Veillonella atypica*
	sat-4A	*Faecalicatena contorta*
	fusC	*Enterococcus* spp. *GMD1E*
	mphC	*Veillonella atypica*
209	*blaZ*	*[Propinobacterium] namnetense SK182B-JCVI*
	mecA	*Macrococcus canis*
	qacA/B	*Bacillus ndiopicus*
	msrA	*Veillonella atypica*
	fusB	*Enterococcus thailandicus*
372	*blaZ*	*[Propinobacterium] namnetense SK182B-JCVI*
	mecA	*Macrococcus canis*
385	*mecA*	*Macrococcus canis*
	blaZ	*[Propinobacterium] namnetense SK182B-JCVI*
	qacA/B	*Bacillus ndiopicus*
386	*mecA*	*Macrococcus canis*
	blaZ	*[Propinobacterium] namnetense SK182B-JCVI*
	qacA/B	*Bacillus ndiopicus*

387	*mecA*	*Macrococcus canis*
		[Propinobacterium] namnetense SK182B-
	blaZ	*JCVI*
	qacA/B	*Bacillus ndiopicus*
	msrA	*Veillonella atypica*
	mphC	*Veillonella atypica*
	fusC	*Enterococcus* spp. *GMD1E*
479	*lnuA*	*Lactobacillus johnsonii*
	mecA	*Macrococcus canis*
		[Propinobacterium] namnetense SK182B-
	blaZ	*JCVI*
	qacA/B	*Bacillus ndiopicus*
	ermC	*Neisseria meningitidis MC58*
	tet(K)	*Streptomyces cinnamoneus*
	fusB	*Enterococcus thailandicus*
620	*mecA*	*Macrococcus canis*
		[Propinobacterium] namnetense SK182B-
	blaZ	*JCVI*
623	*mecA*	*Macrococcus canis*
		[Propinobacterium] namnetense SK182B-
	blaZ	*JCVI*

Table 6.4: Horizontally transferred antibiotic resistance genes in *mecA*$^{+}$ *S. hominis* isolates from East and West London and the predicted donor organisms

5 virulence genes out of 12 were predicted to be horizontally transferred but only in 4 of the 10 isolates (207, 209, 385, 479) (Table 6.5). Interestingly, in isolate 479 among three intercellular adhesins genes responsible for biofilm production, only the *icaA* was predicted to be horizontally transferred but not *icaB* and *icaC*. The majority of the capsule genes that were predicted to be horizontally transferred were found to be donated from *Listeria grayi DSM 20601* and *Nosocomiicoccus* genus.

S. hominis ID	Virulence genes	Closest predicted donors
207	*capsule*	*Nosocomiicoccus* spp. *HMSC067E10*
	capsule	*Nosocomiicoccus ampullae*
	capsule	*Nosocomiicoccus* spp. *HMSC09A07*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	Polysaccharide	
	capsule	*Ornithinibacillus californiensis*
	Polysaccharide	
	capsule	*Vulcanibacillus modesticaldus*
	wbtE	*Amphibacillus sediminis NBRC 103570*
209	*capsule*	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Nosocomiicoccus* spp. *HMSC09A07*
	capsule	*Nosocomiicoccus ampullae*
	capsule	*Nosocomiicoccus* spp. *HMSC067E10*
	Polysaccharide	
	capsule	*Vulcanibacillus modesticaldus*
	Polysaccharide	
	capsule	*Ornithinibacillus californiensis*
	wbtE	*Amphibacillus sediminis NBRC 103570*
385	*icaA*	*Macrococcus caseolyticus JCSC5402*
479	*capsule*	*Salinicoccus luteus DSM 17002*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Nosocomiicoccus* spp. *HMSC059G07*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Nosocomiicoccus ampullae*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	capsule	*Listeria grayi DSM 20601*
	cylR2	*Clostridium* spp. *W14A*

Table 6.5: Horizontally transferred virulence genes in *mecA*$^+$ *S. hominis* isolates recovered from public areas in East and West London

6.3.3 Antibiotic resistance genes carried within a genomic island of *mecA*+ *S. hominis* recovered from public settings in East and West London

Genomic islands were predicted in 10 *mecA*$^+$ *S. hominis* isolates recovered from East and West London based on the Islandviewer 4 method (Table 6.6). The *mecA*$^+$ gene was not identified as being within these predicted genomic islands. From the 10 *mecA*$^+$ *S. hominis*

isolates 2 (20%) had 4 antibiotic resistance genes, 3 (30%) had 2 antibiotic resistance genes, and 2 (20%) had one antibiotic resistance gene and 3 (30%) had no antibiotic resistance genes found within their genomic island. The most common antibiotic resistance genes found within the genomic island were *qacA/B* (n=5); then *blaZ*, *msrA* (n=2); *APH(3')-IIIa*, *sat4A*, *fusB*, *lnuA* and *ANT(4')-lb* (n=1). In addition, 2 isolates had virulence genes predicted within their genomic island. These genes were carried in isolate 385 (*icaA*, *icaB* and *icaC* gene) and in isolate 479 (*cylR2* gene).

***S. hominis* ID.**	**No. of Predicted genomic island regions**	**Antibiotic resistance genes**	**Virulence genes**
207	6	*msrA* *qacA/B*	
208	10	*fusC* *APH(3')-IIIa* *blaZ* *Sat-4A*	
209	7	*msrA* *qacA/B*	
372	7		
385	9	*qacA/B*	*icaA* *icaB* *icaC*
386	8	*qacA/B*	
387	8	*fusC* *blaZ*	
479	6	*fusB* *qacA/B* *ANT(4')-lb* *lnuA*	*cylR2*
620	6		
623	7		

Table 6.6: Genomic islands of 10 *mecA*$^{+}$ *S. hominis* isolates recovered from East and West London

6.3.4 Antibiotic resistance genes carried on plasmids

All *mecA*$^{+}$ *S. hominis* isolates from public settings from East and West London (n=10) were predicted to possess plasmids of which 4 isolates were identified to carry the majority of their antibiotic resistance genes on plasmids (Table 6.7). Between 1 and 5 plasmids were identified per isolate of which 1 isolate (10%) had 6 antibiotic resistance genes carried on plasmids; 2 isolates (20%) had 5 antibiotic resistance gene carried on its plasmids; 1 isolate (10%) had 3 antibiotic resistance genes carried on a plasmid; 3 isolates (30%) had 1 antibiotic resistance gene carried on a plasmid whereas 3 isolates (30%) had no antibiotic resistance gene. Isolate 207 had 5 antibiotic resistance genes encoded on a single genome. The most common antibiotic resistance gene encoded on the plasmid was *qacA/B* (n=6) followed by *fusB* and *ANT(4')-Ib* (n=3), *blaZ*, *mphC* , *ermC* and then *msrA* (n=2) and *lnuA* and *tet(K)*(n=1). A virulence gene *cylR2* was predicted to be carried on a plasmid in isolate 479.

S. hominis ID	Number of plasmids	Chromosome: plasmid-encoded antibiotic genes	Antibiotic resistance genes (and plasmids they were found on)
207	2	1:5	*qacA/B* (Plasmid 1) *ANT(4')-lb* (Plasmid 1) *fusB* (Plasmid 1) *blaZ* (Plasmid 1) *msrA* (Plasmid 1)
208	5	6:1	*msrA* (plasmid1)
209	3	1:6	*qacA/B* (Plasmid 1) *ANT(4')-lb* (Plasmid 1) *blaZ* (Plasmid 1) *fusB* (Plasmid 1) *ermC* (Plasmid 2)
372	1	2:0	No genes
385	1	1:1	*qacA/B* (Plasmid 1)
386	2	1:1	*qacA/B* (Plasmid 1)
387	2	3:3	*qacA/B* (Plasmid 1) *mphC* (Plasmid 1) *msrA* (Plasmid 1)
479	3	2:6	*qacA/B* (Plasmid 1) *ANT(4')-lb* (Plasmid 1) *fusB* (Plasmid 1) *tet(K)* (Plasmid 1) *lnuA* (Plasmid 2) *ErmC* (Plasmid 3)
620	1	2:0	no genes
623	1	2:0	no genes

Table 6.7: **Number of plasmids and the antibiotic resistance genes carried on a plasmid in *mecA*$^{+}$ *S. hominis* isolates recovered from public areas from East and West London**

6.3.5 Phage prediction in *S. hominis* isolates

Phage insertion was predicted in all the *S. hominis* isolates to different degrees of completeness using the PHASTER software (Table 6.8). Two isolates (20%) had complete phage insertion, 1 isolate has questionable phage insertion (if the phage is intact or not due to missing some phage genes) and the rest were incomplete. Complete phage insertion regions were predicted to be phages previous associated with infecting *Staphylococcus*. A few of the incomplete phage insertion regions were predicted to have

homology to phage proteins (integrases, head protein and transposase) that were associated with infection of *Bacillus*, *Prochlorococcus* and *lactobacillus.* One isolate (10%) had 7 phage insertion regions, 4 isolates (40%) had 4 phage insertion regions, 2 isolates (20%) had 3 phage insertion regions, 1 isolate (10%) had 2 phage insertion regions, and 3 isolates (30%) had 1 phage insertion region. No antibiotic resistance gene or virulence genes were predicted to be carried on the phage insertion region.

***S. hominis* ID**	**Most common phages**	**Completeness**
207	*Staphylococcus phage StB12*	Intact
	Staphylococcus phage SPbeta-like	Incomplete
208	*Staphylococcus phage SPbeta-like*	Incomplete
	Staphylococcus phage PT1028	Questionable
209	*Staphylococcus phage StB12*	Intact
	Staphylococcus phage SPbeta-like	Incomplete
372	*Staphylococcus phage SPbeta-like*	Incomplete
385	*Bacillus phage JL*	Incomplete
	Lactobacillus prophage Lj965	Incomplete
	Prochlorococcus phage P-SSM7	Incomplete
	Staphylococcus phage phiIPLA-RODI	Incomplete
	Staphylococcus phage SPbeta-like	Incomplete
	Staphylococcus phage PT1028	Incomplete
	Staphylococcus phage PT1028	Incomplete
386	*Staphylococcus phage SPbeta-like*	Incomplete
387	*Staphylococcus phage SPbeta-like*	Incomplete
479	*Staphylococcus phage phiRS7*	Incomplete
620	*Staphylococcus phage SPbeta-like*	Incomplete
623	*Staphylococcus phage SPbeta-like*	Incomplete

Table 6.8: Phage prediction in *mecA*$^{+}$ *S. hominis* isolates recovered from East and West London; identification of antibiotic resistance genes carried by phages

6.3.6 Phylogenetic analysis of *mecA*$^{+}$ *S. hominis* isolates recovered from East and West London compared with *S. hominis* reference isolates from ENA database

10 *mecA*$^{+}$ *S. hominis* isolates from public areas in East and West London in this study were phylogenetically compared with *S. hominis* isolates from the ENA database that was recovered from different sources to determine their relatedness. This included isolates recovered from clinical samples (from blood; n=11); healthy human skin (n=6); livestock (cows) (n=11) and mosquitos isolates (n=3); isolates recovered from environments (ancient permafrost and air sample from residential area) (n=2) and plant isolates (Kefir seeds and rice seeds) (n=2). SNP core phylogenetic tree of *S. hominis* isolates contained 2 distinct clades (Figure 6.2). Clade A consisted of isolates from the ENA database recovered from livestock (cows), healthy human skin, air in residential areas, whereas clade B consisted of isolates from the ENA database recovered from clinical samples, healthy human skin, Kefir grain and rice seeds, mosquitos and ancient permafrost. Among *mecA*$^{+}$ *S. hominis*, isolates recovered from public areas in East and West London, isolate 385 from public areas in hospitals was found in clade A, whereas the remaining of the isolates were found in clade B. Isolates recovered from hospitals in East and West London (387, 386, 620, 623 and 372) were found in the same subclade and were genetically similar to isolates from the ENA database recovered from healthy humans (ZBW5). Isolates from West London public areas in the community (207, 208 and 209) were on the same subclade and were genetically related to healthy human isolate from skin (UMB022), environmental isolates from ancient permafrost in Russia (MMP2) and Asian Malaria Mosquito bodies (AS1, AS2 and AS3) and Kefir seeds (KR) (Hughes et al., 2016; Kashuba et al., 2017). However, isolate 207, 208 and 209 in this subclade had some divergence in their most common ancestor in comparison to the ENA database isolate. East London isolate 385 was genetically related to isolates recovered from healthy

humans (Hudgins) and air samples from residential areas (H69). Interestingly, isolate 479 recovered from West London hospital was not genetically related to other isolates. All but the mosquito isolates and the permafrost isolates were shown to harbour the *mecA* gene, which was genetically related to isolates from public settings in London. Isolates from the ENA database that were recovered from clinical samples and 8 out of 10 isolates from the ENA database that were recovered from livestock (SNUC 2444, SNUC 5746, SNUC 3403, SNUC 5852, SNUC 4474, SNUC 2620, SNUC 5336 and SNUC 3870) were not genetically related to isolates recovered from other areas.

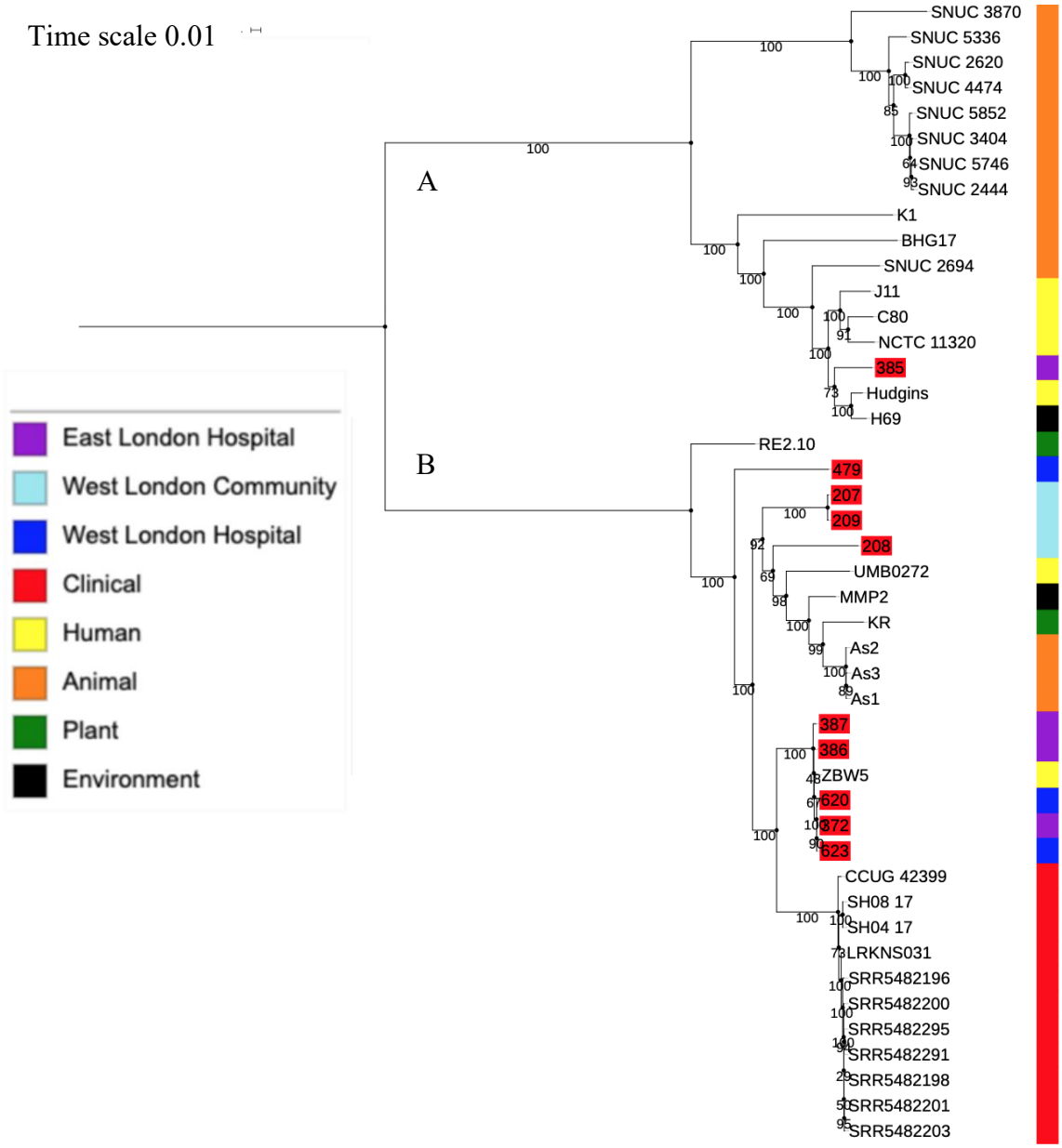


Figure 6.2: **SNP core maximum likelihood phylogenetic tree of 45 *S. hominis* isolates from different sources.** Red highlight= Isolates from this study

PopPUNK analyses identified 23 combined clusters, of which at least 5 distinct clusters were identified for an accessory genome (Figure 6.3). Clinical isolates accessory genomes were found to be clustered together (cluster 5) separately from other isolates recovered from different sources. Additionally, 8 of the *S. hominis* isolates from the ENA database recovered from livestock (SNUC 2444, SNUC 5746, SNUC 3403, SNUC 5852, SNUC

4474, SNUC 2620, SNUC 5336 and SNUC 3870) were found in a different cluster (cluster 2) to isolates recovered from other sources. All of the *mecA*$^{+}$ isolate recovered from public settings in hospitals except for isolate 385 were found in the same accessory genome cluster (cluster 4). This cluster includes isolate 479 which was previously shown not to be phylogenetically related to other isolates by its core genome. These isolates from public settings in cluster 4 were related by their accessory genome to isolates from the ENA database that were recovered from healthy human skin, rice seed (RE2.10) and air samples from residential areas (H69).

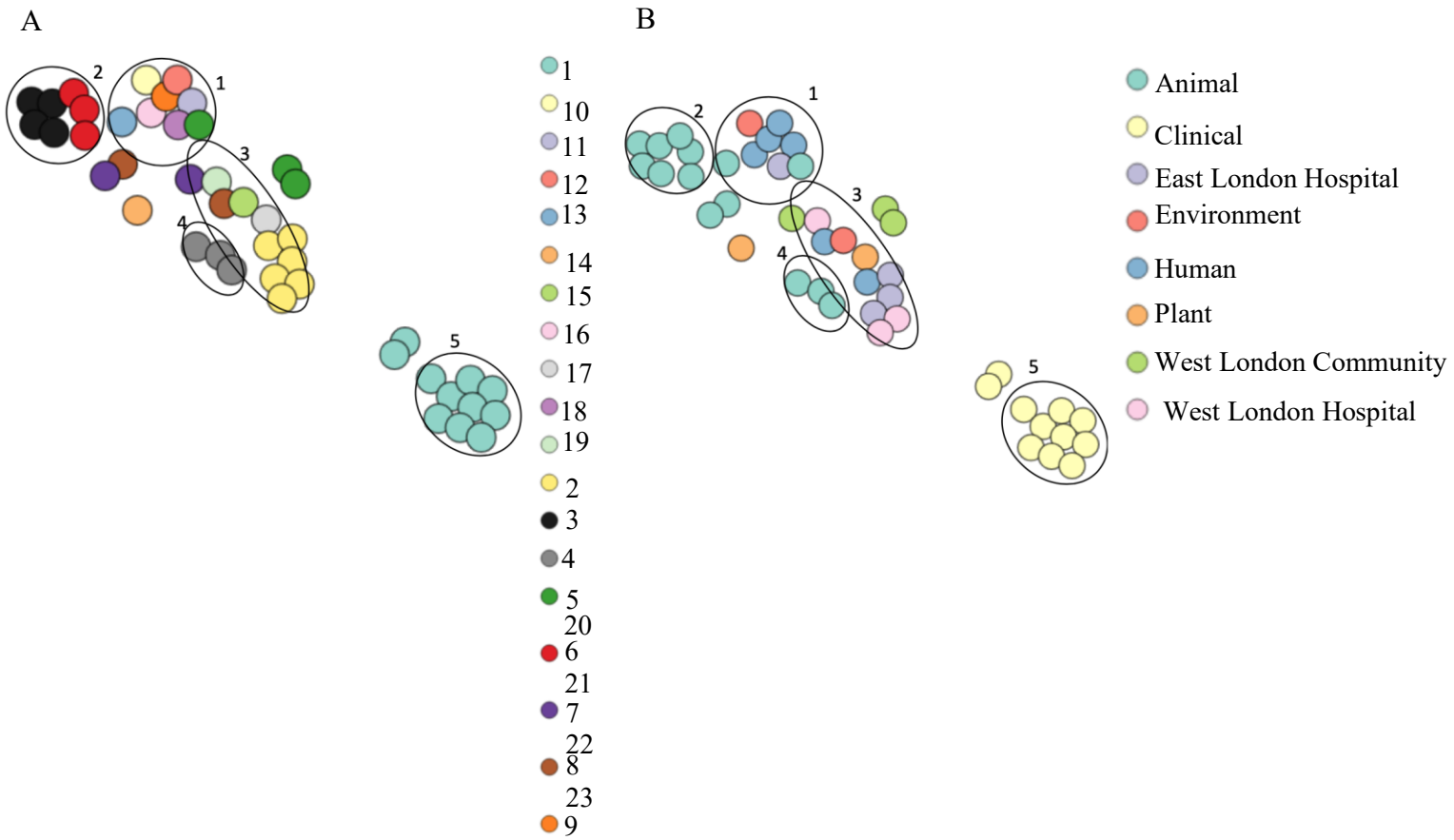


Figure 6 3: **t-SNE analyses of the distance of the accessory genome in 45 *S. hominis* isolates, including *mecA*$^{+}$ *S. hominis* isolates recovered from public settings in East and West London.** The analysis was performed using PopPUNK pipeline**.** Maximum number of mixture components was set at 5 and for the perplexity of the t-SNE set at 10. (A) Combined cluster from PopPUNK analyses; (B) isolation source

6.3.7 Pangenome analysis of *S. hominis*

A pangenome analysis was performed for 45 *S. hominis*, including the 10 *mecA*$^{+}$ *S. hominis* isolates recovered from public settings in East and West London and those obtained from the ENA database. In total, there were 6,386 genes, of which 1,148 (18.0%) were considered to be core. 405 (6.3%) of the genes were softcore gene (present in 95-99% of the isolates); 1,195 (18.7%) of the genes were shell genes (present in 15-95% of the isolates) and 3,638 (57.0%) of the genes were considered cloud genes (present in ≤ 15% of the isolates). A hierarchy clustering heatmap of the 'presence/absence' of the 6,386 genes showed that there was a clustering of isolates obtained from the ENA database (those recovered from clinical samples and 8 out 10 isolates recovered from livestock) (Figure 6.4). Twenty-two unique genes were identified in all isolates from the ENA database recovered from clinical samples, of which 15 genes were predicted as hypothetical. The genes which were not hypothetical were chromosome recombinase genes *ccrA3/B3*; transposition regulatory protein allele *tnpB*; cadmium resistance genes *cadA*, *cadX and cadD* allele; putative DNA repair protein, *radC* and copper-sensing transcriptional repressor *ricR*. For the isolates from the ENA database recovered from livestock, there were no unique genes that were present in all 10 isolates.

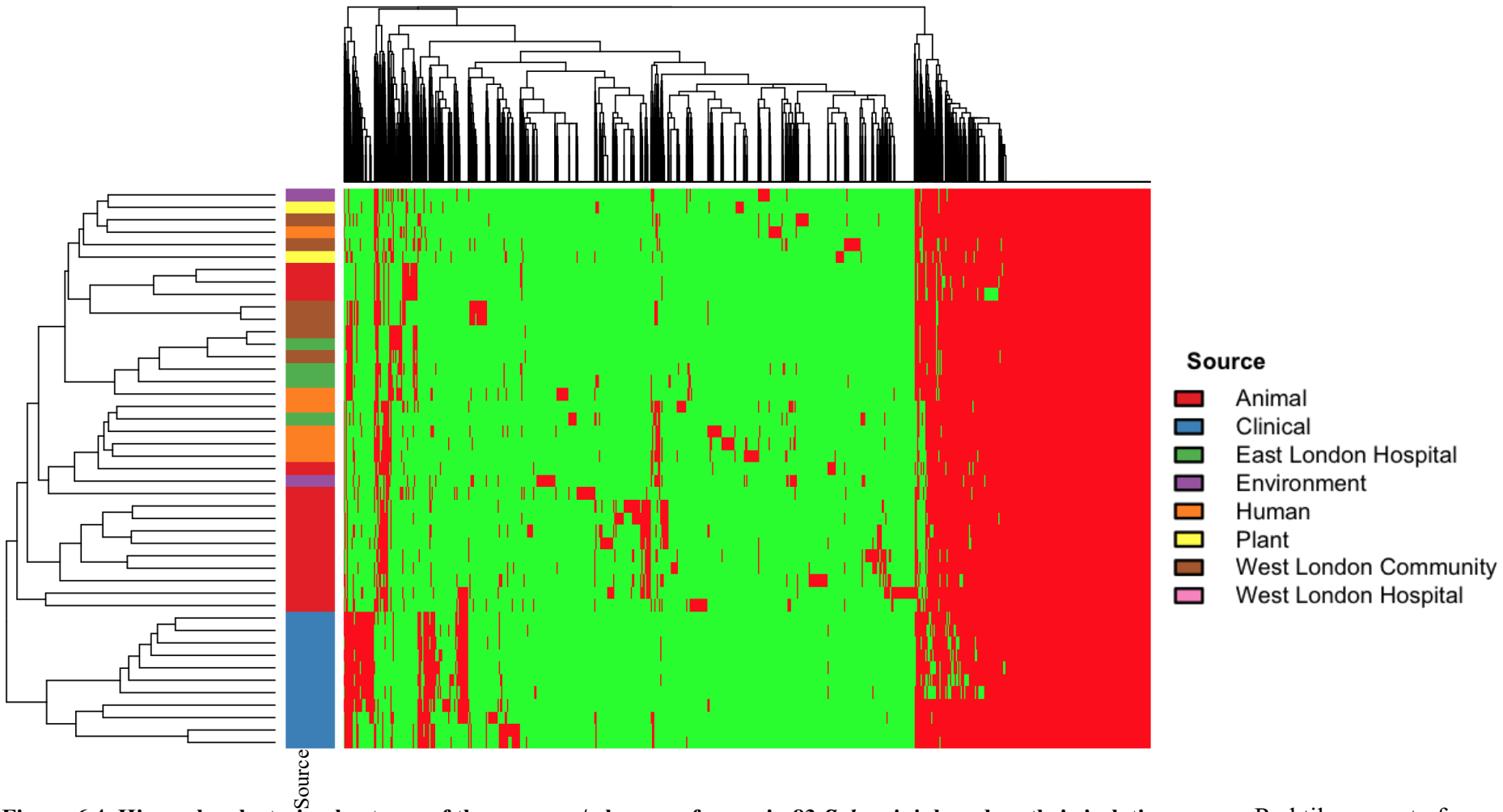


Figure 6.4: **Hierarchy clustering heatmap of the presence/ absence of genes in 83 *S. hominis* based on their isolation source.** Red tile present of gene; green tile absent of gene.

18 antibiotic resistance genes and 16 virulence genes were predicted in 45 *S. hominis* isolates in the pangenome analysis (Table 6.9 and 6.10). No antibiotic resistance genes or virulence genes were ubiquitously found in the 45 *S. hominis*. The most common antibiotic resistance genes predicted were *mecA*, *qacA/B* and *blaZ* (51.1%, 51.1% and 48.9% respectively) and the most common virulence genes predicted were *lip*, *capB* and polysaccharide capsule (97.8%, 95.6% and 93.3% respectively). Antibiotic resistance genes *aad(6)* and *TEM-116* were identified in a single isolate from ancient permafrost (MMP2) and mosquito (As3) respectively. Antibiotic resistance gene *dfrC* was only found in isolates recovered from clinical samples. Interestingly, the intercellular adhesion genes *icaA*, *icaB* and *icaC* were only found in sample 385 from the East London Hospital. From these isolates, 1 isolate (2.2%) had 15 polysaccharide capsule genes; 2 isolates (4.4%) had 14 polysaccharide capsule gene; 6 isolates (13.3%) had 7 polysaccharide capsule genes; 2 isolates (4.4%) had 6 polysaccharide capsule genes; 1 isolate (2.2%) had 3 polysaccharide capsule genes and 29 (64.5%) had 1 polysaccharide capsule genes (Table 6.11). From the *mecA*$^{+}$ *S. hominis* isolates from public settings in East and West London 1 (10%) isolate has 15 polysaccharide capsule gene, 2 (20%) isolate have 6 polysaccharide capsule genes and 7 (70%) isolate has 1 (10%) polysaccharide capsule genes. A hierarchy clustering heatmap shows no clustering based on the source of isolations for antibiotic resistance or virulence genes though isolates from West London community had similar resistance profiles with reference isolates recovered from clinical samples (Figure 6.5A and 6.5B).

Antibiotic resistant genes	List of antibiotic classes	Percentage (%)
TEM-116	cephalosporin, monobactam, penam	2.2
ANT(4')-Ib	aminoglycoside	15.6
qacA/B	fluoroquinolone	51.1
mphC	macrolide	26.7
lnuA	lincosamide	11.1
tet(k)	Tetracycline	35.6
blaZ	Penam	48.9
cat	Phenicol	11.1
msrA	streptogramin, macrolide	33.3
ermC	streptogramin, lincosamide, macrolide	11.1
AAC(6')-le-APH(2")-la	aminoglycoside	17.8
APH(3')-IIIa	aminoglycoside	13.3
mecA	carbapenem, cephalosporin, penam, cephamycin, monobactam	51.1
fusC	fusidic acid	11.1
dfrC	diaminopyrimidine	4.4
sat-4A	nucleoside antibiotic	11.1
fusB	fusidic acid	11.1
aad(6)	aminoglycoside	2.2

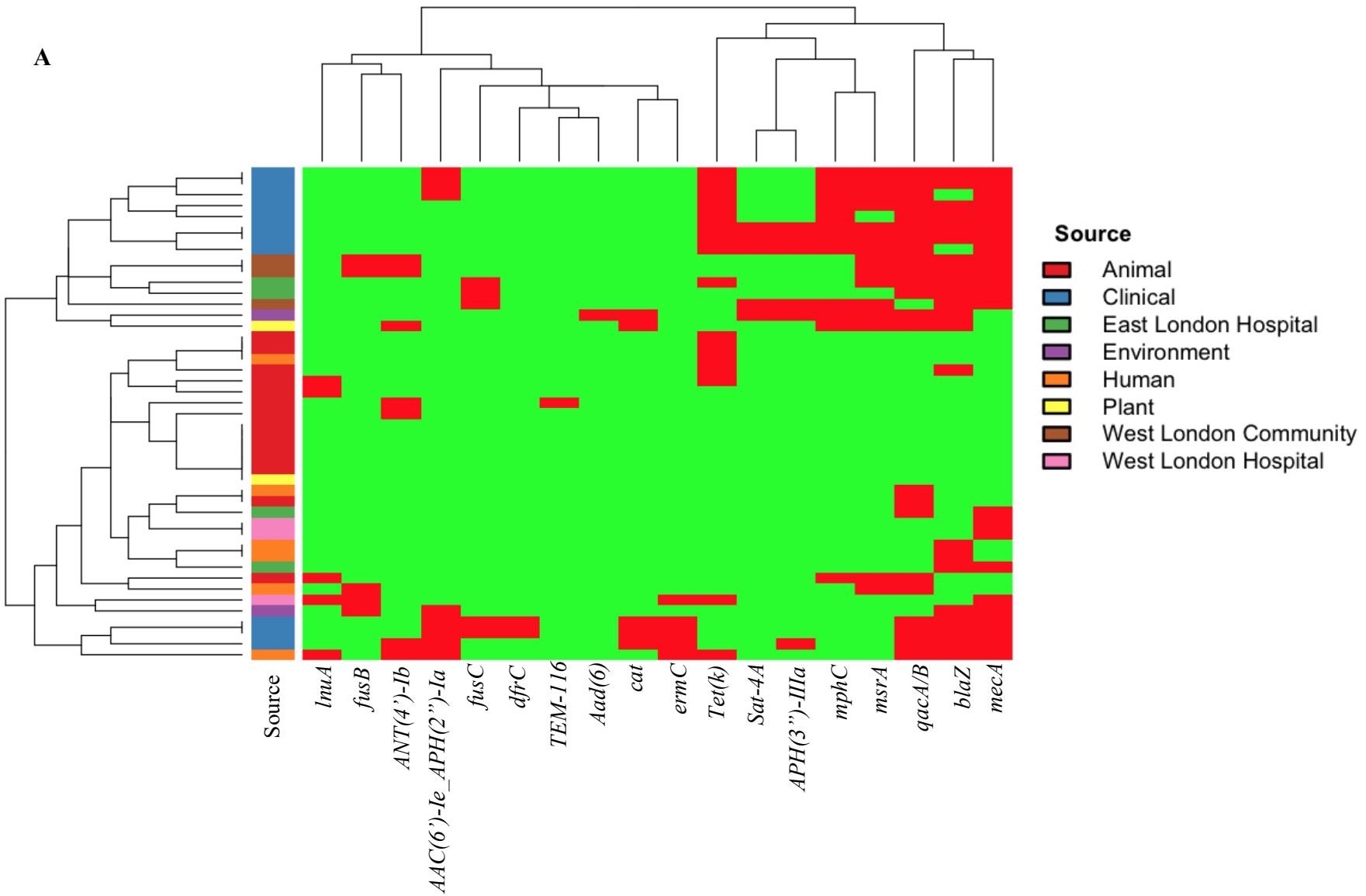
Table 6.9: **Percentage of antibiotic resistance genes in 45 *S. hominis* included in the pangenome analysis**

Virulence genes	Functions	Percentage (%)
atl	Autolysin	86.7
lip	Lipase	97.8
nuc	Thermonuclease	91.1
capsule	Polysaccharide capsule	82.2
capB	Polyglutamic acid capsule	95.6
capC	Polyglutamic acid capsule	84.4
gtaB	Polysaccharide capsule	13.3
capsule (Acinetobacter)	capsule	13.3
uge	Capsule (*Klebsiella*)	8.9
wbtE	LPS	8.9
wbtP	LPS	6.7
cylR2	Cytolysin	4.4
polysaccharide capsule	polysaccharide capsule	93.3
icaA	Intercellular adhesion	2.2
icaB	Intercellular adhesion	2.2
icaC	Intercellular adhesion	2.2

Table 6.10: Percentage of virulence genes in 45 *S. hominis* included in the pangenome analysis

S. hominis ID	Source	Number of capsule genes
As1	Animal	1
As2	Animal	1
As3	Animal	0
BHG17	Animal	1
K1	Animal	3
SNUC_2444	Animal	1
SNUC_2620	Animal	7
SNUC_2694	Animal	7
SNUC_3404	Animal	1
SNUC_3870	Animal	14
SNUC_4474	Animal	1
SNUC_5336	Animal	14
SNUC_5746	Animal	1
SNUC_5852	Animal	1
CCUG_42399	Clinical	6
SH04_17	Clinical	1
SH08_17	Clinical	1
SRR5482191	Clinical	1
SRR5482195	Clinical	0
SRR5482196	Clinical	1
SRR5482198	Clinical	1
SRR5482200	Clinical	0
SRR5482201	Clinical	1
SRR5482203	Clinical	1
LRKNS031	Clinical	1
372	East London Hospital	1
385	East London Hospital	1
386	East London Hospital	1
387	East London Hospital	1
H69	Environment	7
MMP2	Environment	1
C80	Human	7
HUDGINS	Human	1
J11	Human	7
NCTC_11320	Human	7
UMB0272	Human	1
ZBW5	Human	1
RE2.10	Plant	1
KR	Plant	1
207	West London Community	6
208	West London Community	1
209	West London Community	6
479	West London Hospital	15
620	West London Hospital	1
623	West London Hospital	1

Table 6.11: Number of capsule genes to isolation source in the 45 *S. hominis* isolates. Blue highlight= East London Hospital. Pink highlight= West London community and green highlight=West London Hospital.


A
Source
lnuA
fusB
ANT(4')-Ib
AAC(6')-Ie_APH(2'')-Ia
fusC
dfrC
TEM-116
Aad(6)
cat
ermC
Tet(k)
Sat-4A
APH(3'')-IIIa
mphC
msrA
qacA/B
blaZ
mecA
Source
Animal
Clinical
East London Hospital
Environment
Human
Plant
West London Community
West London Hospital

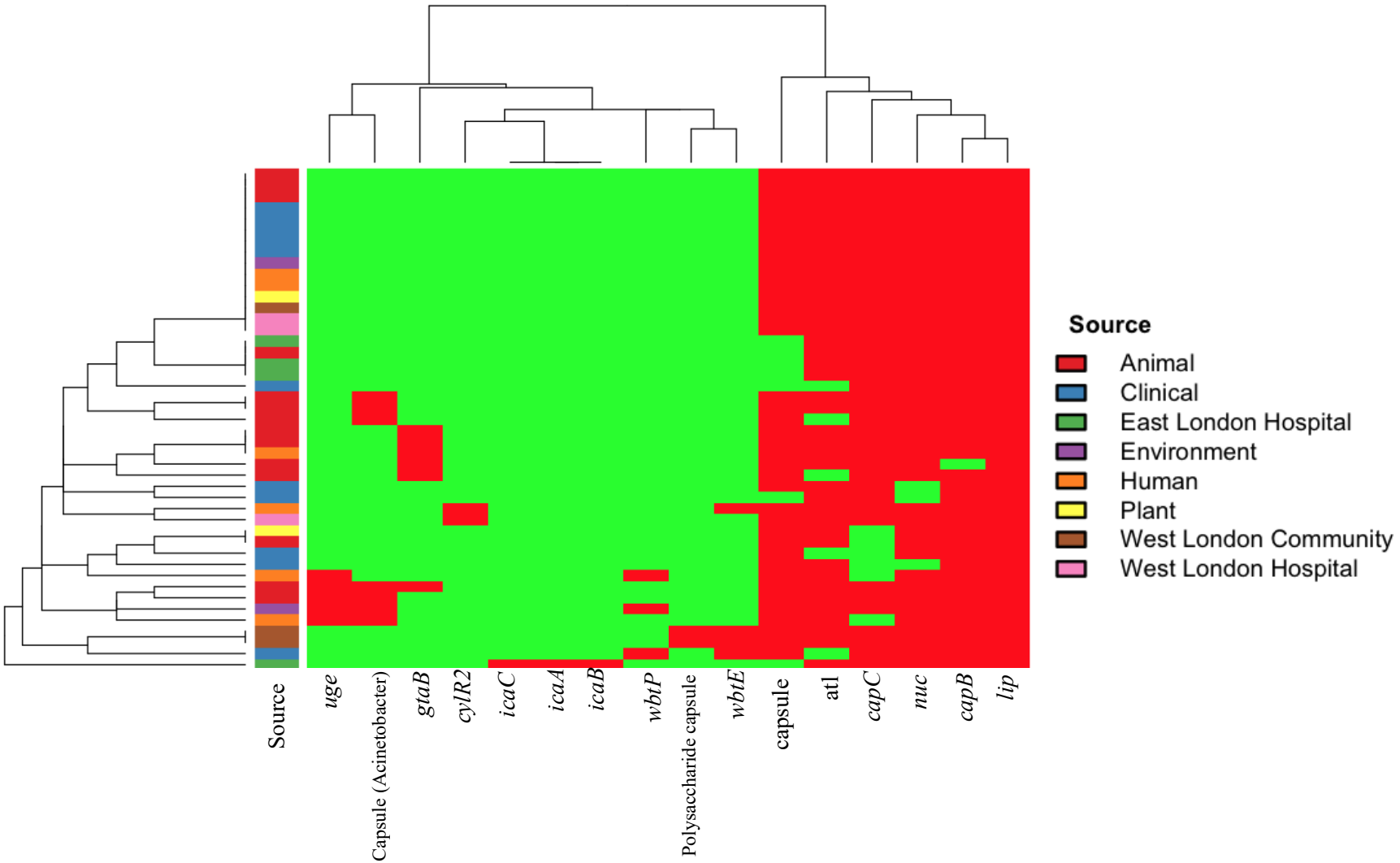


Figure 6.5: Hierarchy clustering heatmap of antibiotic resistance and virulence genes found in 45 *S. hominis* isolates used for the pangenome analysis.(A) Hierarchy clustered heatmap of antibiotic resistance genes by Isolation source; (B) Hierarchy clustered heatmap of virulence genes by isolation source. Red tile present; green tile absent of gene.

The COG family group was identified for the core and accessory genes in *S. hominis* pangenome analysis (Table 6.12). In the core genome, the highest portion of the genes was predicted to be the general function prediction only (12.5%) followed by amino acid transport and metabolism (10.0%) and translation, ribosomal structure and biogenesis (9.8%). In the accessory genome, the highest portion of the genes was predicted to be the general function prediction only (15.9%) followed by function unknown prediction only (13.0%) and transcription (10.7%).

	Percentage (%) of genes found in the pangenome	
Function	**Core**	**Accessory**
Translation, ribosomal structure and biogenesis	9.5	2.5
RNA processing and modification	0.0	0.0
Transcription;	6.3	9.4
Replication, recombination and repair	4.8	15.0
Chromatin structure and dynamics	0.1	1.2
Cell cycle control, cell division, chromosome partitioning	0.9	3.1
Nuclear structure	0.0	0.0
Defence mechanisms	1.2	2.2
Signal transduction mechanisms	3.2	8.5
Cell wall/membrane/envelope biogenesis	3.4	0.1
Cell motility	0.2	0.0
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	1.2	0.9
Posttranslational modification, protein turnover, chaperones	3.7	1.6
Energy production and conversion	7.4	2.8
Carbohydrate transport and metabolism	5.8	8.0
Amino acid transport and metabolism	9.5	6.6
Nucleotide transport and metabolism	4.6	1.7
Coenzyme transport and metabolism	6.2	2.4
Lipid transport and metabolism	2.7	2.3
Inorganic ion transport and metabolism	6.6	5.7
Secondary metabolites biosynthesis, transport and catabolism	1.5	1.6
General function prediction only	11.8	13.6
Function unknown	9.7	10.7

Table 6.12: COG family group of genes in the core and accessory genome of *S. hominis* isolates

The COG family group was identified for the unique genes in *mecA*$^{+}$ *S. hominis* isolates from public settings in East and West London (Table 6.13). 334 genes were identified as being unique to public setting isolates from East and West London. The highest portion of the unique genes belonged to replication, recombination and repair (19.4%); general function prediction only (16.3%) and function unknown (13.4%).

Function	Percentage (%) of genes in the pangenome unique to *S. hominis* isolates from East and West London
Translation, ribosomal structure and biogenesis	1.7
Transcription	13.2
Replication, recombination and repair	19.4
Cell cycle control, cell division, chromosome partitioning	0.5
Defense mechanisms	6.7
Signal transduction mechanisms	1.3
Cell wall/membrane/envelope biogenesis	8.2
Posttranslational modification, protein turnover, chaperones	0.6
Carbohydrate transport and metabolism	4.7
Amino acid transport and metabolism	6.5
Nucleotide transport and metabolism	0.9
Coenzyme transport and metabolism	0.9
Lipid transport and metabolism	0.7
Inorganic ion transport and metabolism	5.0
General function prediction only	16.3
Function unknown	13.4

Table 6.13: **Unique genes of COG family in 10 *mecA*$^{+}$ *S. hominis* isolates recovered from public setting in East and West London**

Comparative analysis of the accessory genome in *mecA*$^+$ *S. hominis* isolates from public settings in East and West London showed that the East London *mecA*$^+$ isolates had 320 genes that were not found in West London isolates, whereas isolates recovered from West London had 517 genes that were not found in *mecA*$^+$ East London isolates. Only 1 gene was ubiquitous in *mecA*$^+$ East London isolates that was not found in *mecA*$^+$ West London isolates. This was the *arsC* gene that encodes for arsenate reductase (glutaredoxin), whereas there were no genes that were ubiquitous to *mecA*$^+$ West London isolates but not present in East London isolates. General Public setting isolates were predicted to have 311 unique genes not identified in *mecA*$^+$ isolates in public areas in hospitals, whereas *mecA*$^+$ isolates recovered from public areas in hospital had 625 unique genes not found in public areas in the community. 6 genes were ubiquitous to isolates recovered from general public settings, but these were not found in isolates recovered from public areas in hospitals, whereas 3 genes were ubiquitous to *mecA*$^+$ isolates from public areas from the hospital, which were not present in isolates from general public settings. The known function of unique ubiquitous genes in *mecA*$^+$ isolates recovered from public areas from the community was bacteriophage integrase; spore coat protein and recombinase/resolvase. Whereas, the known function of the ubiquitous genes unique to isolates recovered from *mecA*$^+$ isolates from public areas in hospitals was oligopeptide; ABC superfamily, ATP binding cassette transporter and membrane protein. There was no difference in the COG function of the genes that were found in *mecA*$^+$ isolates recovered from East and West London. There was a higher proportion of unique genes of COG function in *mecA*$^+$ isolates recovered from general public settings than in the *mecA*$^+$ isolates recovered from public areas in hospitals that had a function for transcription (12.6% and 8.0% respectively (P=0.0255)); replication, recombination and repair (21.9% and 15.6% respectively (P=0.0411)) defense mechanisms (12.6% and 3.5%) respectively

(P=<0.0001)) and translation, ribosomal structure and biogenesis (3.9% and 1.3% respectively (P=0.0104)) (Table 6.14). There was a higher proportion of unique genes of the COG function in *mecA*$^{+}$ isolates recovered from public areas in hospitals compared with *mecA*$^{+}$ isolates recovered from general public settings with the function of inorganic ion transport and metabolism 11.9% and 3.4% respectively (P=<0.0001); energy production and conversion 3.3% and 0.9% respectively (P=0.0265); amino acid transport and metabolism 6.1% and 2.6% respectively (P=0.0216) (Table 6.15).

	Percentage (%) of unique genes in the pangenome from East and West London	
COG Function	**East London**	**West London**
Translation, ribosomal structure and biogenesis	1.3	1.2
RNA processing and modification	0.0	0.0
Transcription	8.0	8.3
Replication, recombination and repair	15.6	17.5
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	0.8	0.3
Nuclear structure	0.0	0.0
Defence mechanisms	3.4	3.6
Signal transduction mechanisms	2.0	3.1
Cell wall/membrane/envelope biogenesis;	10.7	10.8
Cell motility	0.3	0.6
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	1.0	0.6
Posttranslational modification, protein turnover, chaperones	0.9	1.0
Energy production and conversion	3.3	1.9
Carbohydrate transport and metabolism	7.5	8.1
Amino acid transport and metabolism	6.1	6.3
Nucleotide transport and metabolism	0.6	1.2
Coenzyme transport and metabolism	1.5	0.7
Lipid transport and metabolism	1.4	1.9
Inorganic ion transport and metabolism	11.9	9.1
Secondary metabolites biosynthesis, transport and catabolism	2.1	3.3
General function prediction only	10.4	10.4
Function unknown	11.1	10.2

Table 6.14: **Difference in COG function of the unique genes in *mecA*$^+$ *S. hominis* isolates recovered from different environmental sites in East and West London.** *= The area where there was significantly higher (p=<0.05) percentage of genes for that COG function

	Percentage (%) of unique genes in the pangenome from London public settings in the Community and public areas in Hospital	
COG function	**Community**	**Hospital**
Translation, ribosomal structure and biogenesis	3.9	1.3
RNA processing and modification	0.0	0.0
Transcription	12.6*	8.0
Replication, recombination and repair	21.0*	15.6
Chromatin structure and dynamics	0.0	0.0
Cell cycle control, cell division, chromosome partitioning	2.2	0.8
Nuclear structure	0.0	0.0
Defence mechanisms	12.6*	3.4
Signal transduction mechanisms	1.3	2.0
Cell wall/membrane/envelope biogenesis;	7.2	10.7
Cell motility	0.0	0.3
Cytoskeleton	0.0	0.0
Extracellular structures	0.0	0.0
Intracellular trafficking, secretion, and vesicular transport	0.0	1.0
Posttranslational modification, protein turnover, chaperones	0.0	0.9
Energy production and conversion	0.9	3.3*
Carbohydrate transport and metabolism	8.7	7.5
Amino acid transport and metabolism	2.6	6.1*
Nucleotide transport and metabolism	0.9	0.6
Coenzyme transport and metabolism	1.8	1.5
Lipid transport and metabolism	0.7	1.4
Inorganic ion transport and metabolism	3.4	11.9*
Secondary metabolites biosynthesis, transport and catabolism	0.7	2.1
General function prediction only	12.5	10.4
Function unknown	7.0	11.1

Table 6.15: Difference in COG function the unique genes in *mecA*$^{+}$ *S. hominis* isolates recovered from different environmental sites in public settings in the community and public areas in hospitals.*= The area where there was significantly higher (p=<0.05) percentage of genes for that COG function.

6.4 Discussion

6.4.1 *S. hominis*: Virulence genes and HGT

Similar to *S. haemolyticus,* S. *hominis* had a smaller pool of virulence genes compared to S. *epidermidis*. There are only a few studies that looked at *S. hominis* virulence genes (Calkins et al., 2016; Saiping Jiang et al., 2012; Szczuka et al., 2018). Among the *mecA*$^+$ isolates recovered from public areas in East and West London, *Bacillus* capsule genes, *capB* and *capC* were found in 95.6% and 84.4% respectively of all isolates used in the pangenome analyses including the *mecA*$^+$ from public settings and the isolates from the ENA database. These virulence genes are normally found in *Bacillus,* but the polyglutamic acid capsule has not been reported before in *S. hominis* (Kocianova et al., 2005). The genes *capB* and *capC* were not identified as being horizontally transferred from another organism into *mecA*$^+$ isolates recovered from East and West London (Calkins et al., 2016). These genes, therefore, may have originated from *S. hominis* or have been transferred from another staphylococcal species as the HGTector pipeline can only distinguish genes transferred from bacteria from a different genus (Zhu et al., 2014). On the other hand, the polysaccharide capsule genes that were associated with *S. hominis* and other staphylococci were predicted to be horizontally transferred (Flahaut et al., 2008; O'Riordan and Lee, 2004). 3 of the *mecA*$^+$ isolates from public settings (207, 209 and 479) that were predicted to have multiple polysaccharide capsule genes that have been horizontally transferred from another genus (Flahaut et al., 2008; O'Riordan and Lee, 2004). For the *mecA*$^+$ *S. haemolyticus* isolates recovered from public settings, the predicted donor genus for these capsule genes was *Listeria* and *Nosocomiicoccus*. Isolates from ENA database and *mecA*$^+$ S. *hominis* from this study either have 1, 6, 7, 14 or 15 genes responsible for encoding for polysaccharide capsule. These finding were similar to the *S. haemolyticus* pangenome analyses in this study. However, there were some

exceptions, for example, *S. hominis* isolates had a higher number of polysaccharide capsule genes in isolates recovered from livestock (SNUC_3870 and SNUC_5336) than in clinical samples and one of the isolates in this study recovered from West London hospital (479) had 15 capsule genes. This would indicate that the polysaccharide capsule has evolved in isolates from livestock rather than in humans for S. *hominis*. Expect that these genes in *S. hominis* require multiple polysaccharide genes to produce virulence factors as it has been shown for *S. aureus* and *S. haemolyticus* to code for different functions (for example chain length determination; putative tyrosine protein kinase, putative phosphotyrosine-protein phosphatase, putative 4,6-dehydratase, glycolysis transferase and aminotransferase) in the production of the polysaccharide capsule (Flahaut et al., 2008; Kuipers et al., 2016) Additionally, a gene was found which encodes for capsules in *Acinetobacter* spp. and *Klebsiella* spp. (*uge* gene). *Uge* gene has been identified as being important in *Klebsiella pneumoniae* capsule production whereas the *Acinetobacter* capsule gene is shown to be one of the genes responsible for immune invasion in *Acinetobacter baumannii* (Regué et al., 2004; Singh et al., 2019). These virulence determents from *Klebsiella pneumoniae* and *Acinetobacter baumannii* may not be encoded in *S. hominis,* but these findings indicate that other virulence genes could potentially horizontally transfer from these species making *S. hominis* a more virulent variant.

Interestingly, an isolate recovered from the East London hospital (385) was identified to possess the genes responsible for *icaA*, *icaB* and *icaC* gene but not *icaD*. These genes were not present in any other *mecA*$^{+}$ *S. hominis* isolates from public settings in East and West London or from the *S. hominis* isolates from the ENA database. This contradicts studies that found the *icaADBC* operon was typically found in *S. hominis* isolates from clinical isolates recovered from blood (Soroush et al., 2017; Szczuka et al., 2015). The

small collection of WGS *S. hominis* available from the ENA database might have biased this data for non-biofilm producing *S. hominis*. It would be expected to see biofilm encoding genes in high abundance in clinical isolates as it is an important virulence factor as well as the *mecA*$^{+}$ isolates from high-frequency touched surfaces as the adaption gives the organism a better chance to survive on abiotic surfaces (Khelissa et al., 2017). It is possible there may be another mechanism that is not known that can also produce biofilm similar to that found in *S. haemolyticus* isolates (Fredheim et al., 2009). *IcaA* gene found in the *mecA*$^{+}$ isolate recovered from East London was predicted to be horizontally transferred but not the *icaB* or *icaC* gene. The predicted donor for *icaA* was *Macrococcus caseolyticus* the same donor predicted in *mecA*$^{+}$ *S. epidermidis* isolates from general public settings. Additionally, these genes were predicted to be located within a genomic island region. This could indicate that biofilm producing genes can be found within a mobile genetic cassette in this *S. hominis* isolate and been transferred from another species.

Similar to *S. haemolyticus* isolates, genes were identified that were responsible for lipopolysaccharide production for host immune invasion in *Francisella* (*wbtE and wbtP*) and the gene that is involved in cytolysin production in *Enterococcus (cylR2).* These genes were found in 8.9%, 6.7% and 4.4% of *mecA*$^{+}$ isolates recovered from public settings and *S. hominis* isolate from the ENA database, respectively. *WbtE* and c*ylR2* genes in *mecA*$^{+}$ *S. hominis* isolates recovered from public setting were identified to be horizontally transferred from *Amphibacillus sediminis NBRC 103570* and *Clostridium* spp. *W14A* respectively. Cytolysin has been previously reported in *Clostridium* species and lipopolysaccharide in *Amphibacillus* species (Antunes et al., 2016; Heuck et al., 2010). In *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings the predicted donor for *cylR2* (*Virgibacillus* spp. *SK37 and Streptococcus* spp. *HMSC10A01*) were

different than in *mecA*$^{+}$ *S. hominis* isolates from public settings which suggest that this gene in *S. haemolyticus* and *S. hominis* may have different origins. As both of these genes were found within an operon with other genes that were responsible for encoding for cytolysin and lipopolysaccharide production, it is expected that the *cylR2* gene alone will encode for cytolysin or that the *wbtE* and *wbtP* would encode for lipopolysaccharide production. These findings suggest that virulence genes may horizontally transfer to *S. hominis* from a different genus making it more virulent (Coburn et al., 2004; Twine et al., 2012).

6.4.2 *S. hominis*: Antibiotic resistance genes and HGT

In the previous chapters, it was identified that many of the antibiotic resistance genes were carried on plasmids in *mecA*$^{+}$ isolates recovered from public settings in East and West London. This was also observed in the *mecA*$^{+}$ *S. hominis* isolates recovered from public settings. In one of the isolates (207), 10 out of 12 antibiotic resistance genes were encoded on a single plasmid. This shows that *S. hominis* isolates recovered from public settings acquire genes via plasmid conjugation or transformation. The uptake of plasmids is a common method of horizontal gene transfer in bacteria which gives them the ability to quickly obtain multiple antibiotic resistance genes from a single plasmid or multiple plasmids making the bacterium multidrug resistance (Millan, 2018). Interestingly, the *blaZ* was identified to be donated from another organism which was not predicted in *mecA*$^{+}$ *S. epidermidis* and *S. hominis* isolates from public settings from East or West London. The predicted donor was *Propinobacterium namnetense* though it most likely from an ancestor of the predicted donor species (Zhu et al., 2014). This suggests that the *blaZ* gene in *S. hominis* was transferred from another genus other than staphylococci. It was determined that antibiotic resistance genes were found within the host genome in 3 of the *mecA*$^{+}$ *S. hominis* isolates recovered from public settings (isolate 208, 209 and 387).

Antibiotic resistance genes found in different genomic island regions were the *blaZ, fusC, msrA*, *APH (3')-IIIa* and *sat4A*. Interestingly, all these genes were identified to be horizontally transferred from another donor except for *APH (3')-IIIa.* This indicates that these genes have been transferred from a different genus of bacteria either via plasmid which has then been integrated into the host genome or via transposons or mobile cassettes. The *APH(3')-IIIa* gene was most likely transferred from another *S. hominis* isolate or other staphylococci isolate as HGTtector pipeline can only detect horizontal gene transfer from bacteria of a different genus to the one that was being analysed (Zhu et al., 2014).

Pangenome analyses showed that antibiotic resistance genes *TEM-116* and *aad(6)* were identified from a single isolate (As1 and MMP2, respectively). Although there is lack of data on antibiotic resistance gene abundance in *S. hominis,* studies have shown that these genes are rarely found in *S. epidermidis* based on the Comprehensive Antibiotic Resistance Database (Jia et al., 2017). Interestingly, these genes were identified in isolates recovered from ancient permafrost (MMP2) for *aad(6)* and mosquitoes bodies (*As3*) for *TEM-116*, which encode for aminoglycoside resistance and broad-spectrum beta-lactamase, respectively. Finding the *TEM-116* in isolate recovered from ancient permafrost would suggest that *TEM-116* evolved before the antibiotic resistance genes were mass-produced and commonly used. This gene protects against fungi that produce beta-lactam antibiotics as a defence mechanism (Gao et al., 2017). Additionally, *dfrC* gene, which encodes resistance for the diaminopyrimidine antibiotics was only found in isolates recovered from clinical samples (blood). These data were similar to findings from this study in *S. haemolyticus* analyses. However, in *S. epidermidis* isolates recovered from different sources that possess the *dfrC* gene was relatively higher (94.4%). This suggests

that *dfrC* gene originated from *S. epidermidis* and has horizontally been transferred to *S. hominis* and *S. haemolyticus* isolates in hospitals.

Pangenome analyses of *S. hominis* isolates showed that these isolates possessed less *mecA* and *blaZ* (51.11%, 48.9%) compared to that of *S. epidermidis* (65.6% and 80% respectively) and *S. haemolyticus* (78.3% and 72.3% respectively) isolates analysed in previous chapters. This may be due to that *S. hominis* is not as frequently isolated as *S. epidermidis* and *S. haemolyticus in* hospitals. In addition, there are fewer studies reporting cases of *S. hominis* infections (Chaves et al., 2005; Spanu et al., 2003). Therefore, *S. hominis* isolates may not have been in frequent contact with beta-lactam antibiotics to gain or maintain *mecA* and *blaZ* antibiotic resistance genes.

6.4.3 HGT of *mecA*$^+$ *S. hominis* isolates recovered from East and West London

The HGT derived genome of the *mecA*$^+$ *S. hominis* recovered from public settings in East and West (7.8% to 9.5%) was similar to *mecA*$^+$ *S. epidermidis* (8.6% to 10.1%) and *mecA*$^+$ *S. haemolyticus* (8.9% to 13.5%) recovered from the same places. Interestingly, in *mecA*$^+$ *S. hominis* isolates the genus that donated the most genes was *Salinicoccus*, whereas in *S. haemolyticus* and *S. epidermidis* this was predicted to be the third most common donor genus. The *Salinicoccus* genus can live in high pH environments and is not associated with humans, animals or other environments (Coburn et al., 2004; Twine et al., 2012). *Salincicoccus* and *Staphylococcus* share a common ancestor and belong to the same *Staphylococcaceae* family.

Intact phages were identified in 2 of the 10 *S. hominis* isolates recovered from public settings (isolate 207 and 209). The other isolates were identified to have phages that were considered incomplete. This would suggest that an intact phage insertion sequence integrated recently into the host and that integrated genes from the phage have not been

deleted due to them not conferring adaption for survival and/or a metabolic burden (Ramisetty and Sudhakari, 2019). The gene which was identified to be intact was *Staphylococcus* phage StB12. This has previously been identified in *mecA*$^{+}$ *S. epidermidis* and *mecA*$^{+}$ *S. haemolyticus* isolates recovered from public settings and been also reported in *S. hominis* and *S. capitis* (Deghorain et al., 2012). This shows that *S. hominis* can acquire virulence genes from other staphylococci. Therefore, it is possible for *S. hominis* to horizontally transfer the *S. aureus* staphylokinases, superantigens, Panton-Valentine leukocidin virulent genes which are found to be transferred to another *S. aureus* isolates via phages (Goerke et al., 2009). Incomplete phages that show homology to phages associated with *Bacillus*, *Synechococcus*, *Prochlorococcus* and *Lactobacillus*. Due to incomplete phage sequence in other *S. epidermidis* and *S. haemolyticus* these phage sequences are phages that only infect staphylococci that have not been characterised yet.

6.4.4 Phylogenetic and Pangenomic analysis of *S. hominis* isolates

Similar to *S. epidermidis* and *S. haemolyticus* SNP core phylogenetic analysis, the *mecA*$^{+}$ *S. hominis* isolates recovered from public settings in East and West London were genetically similar to isolates obtained from the ENA database that have previously been isolated from other sources. Although there was no *mecA*$^{+}$ *S. hominis* isolates from this study that were genetically related to isolates to those obtained from the ENA database that were recovered from clinical samples (blood) or livestock (cows), however they were related to isolates recovered from healthy human skin (Hudgins and ZBW5), air samples from residential areas (H69), from mosquitos' bodies (As1, As2 and As3), ancient permafrost (MMP2) and Kefir seeds (KR) (Hughes et al., 2016; Rivera-Perez et al., 2016). This suggests that mosquitoes could be possible vectors for transmitting *S. hominis* while feeding on their host and that genetically these *mecA*$^{+}$ *S. hominis* isolates recovered from general public settings have has not evolved much since ancient times

(Hughes et al., 2016). Mosquitos are vectors for viruses, protozoa and parasites that can spread and cause disease in humans and animals but currently, it is unreported if they can transfer and initiate bacterial infections (Huntington et al., 2016). Additionally, it is possible that mosquitoes could transfer livestock-associated *S. hominis* to humans. From this study, it was deduced that the isolates recovered from mosquitos is same genetic lineages of isolates originate from humans as they were found within the same subclade of isolate ZBW5 which was recovered from healthy human skin. Additionally, the findings that the *mecA*$^{+}$ *S. hominis* isolates (385) in this study belonged to same genetic lineage as the isolate recovered from an air sample in a residential area (H69) suggests that *S. hominis* can be transmitted through the air from humans to high frequency touched surfaces or vice versa (Lymperopoulou et al., 2017). *S. hominis* isolates recovered from clinical samples from the ENA database were found together in the same subclades of the core phylogenetic tree (Figure 6.2) and clustered together by their accessory genome from the t-SNE analyse (Figure 6.3) and pangenome analyses (Figure 6.4). 8 out of 10 *S. hominis* livestock isolates (SNUC 2444, SNUC 5746, SNUC 3403, SNUC 5852, SNUC 4474, SNUC 2620, SNUC 5336 and SNUC 3870) were found in the same cluster by their core and accessory genomes. This would suggest that *S. hominis* recovered from clinical isolates and the majority of livestock isolates has evolved separately from other sources and have not spread to different niches. In addition, these results suggest that *S. hominis* recovered from clinical samples are nosocomial isolates (Chaves et al., 2005). This is further supported for *S. hominis* by the fact that these clinically associated isolates share 22 genes which were unique to them only. This included a gene which encodes for cadmium and copper resistance. Previous studies have shown the importance of cadmium resistance in *Helicobacter pylori and Listeria monocytogenes* virulence copper resistance is also important for *S. aureus* survival within macrophages (Purves et al., 2018; Stähler

et al., 2006). Currently, it is unknown whether these genes can benefit to the virulence of *S. hominis*. Although from virulence gene profiles (Figure 6.5B) of 83 *S. hominis* isolates from public settings, ENA database animals' isolates, ENA database isolates from healthy humans and ENA database isolates from the natural environment had the same virulence genes found in the ENA database isolates recovered from clinical samples. This suggests that in *S. hominis*, there may be other virulent factors that have not been characterised before. *S. hominis* isolates from public areas in hospitals were found within the same cluster in their accessory genome. This would indicate that these species in these areas have a similar pool of genes that can be horizontally transferred due to the similarities in the bacterial populations in that geographical area or environment required to survive in these niches (Segerman, 2012).

As with the pangenome analysis of *S. epidermidis* and *S. haemolyticus* isolates from previous chapters *S. hominis* was also shown to be an open genome in which it seems there is an unlimited number of genes as part of the species gene pool (Figure 6.4). This suggests that *S. hominis* can be found in different environments within mixed microbial communities in which they can exchange genetic material to increase their gene pool (Bosi et al., 2015). In a similar manner to *mecA*$^+$ *S. epidermidis* and *mecA*$^+$ *S. haemolyticus* isolates from public settings had a high percentage of unique genes in the *mecA*$^+$ *S. hominis* isolates recovered from public settings that belonged to replication, recombination and repair (19.43%), and the COG group. Although this was less than what was found in *mecA*$^+$ *S. epidermidis* (24.41%) and *mecA*$^+$ *S. haemolyticus* (34.25%) isolates from public settings. This may indicate fewer novel genes for these functions in *S. hominis* recovered from public settings. Additionally, there was a higher proportion of genes unique to *mecA*$^+$ *S. hominis* isolates recovered form general public settings that code for replication, recombination and repair (20.97%) COG group compared to *mecA*$^+$

isolates (15.6%) recovered from public areas in hospitals. This was similar to that observed in this study for *mecA*$^{+}$ *S. haemolyticus* isolates recovered from general public settings and public areas in hospitals. This would indicate that these bacterial species from general public settings are exposed to environmental factors such as to UV light, ionizing radiation and genotoxic chemicals which can cause DNA damage compared to those in hospitals (Žgur-Bertok, 2013). Additionally, *mecA*$^{+}$ *S. hominis* isolates recovered from public areas in the community had a higher abundance of unique genes that encode for transcription, defense mechanisms and translation, ribosomal structure and biogenesis. Whereas *mecA*$^{+}$ isolate in public areas in hospitals had a higher abundance of unique genes which encode for inorganic ion transport and metabolism, energy production and conversion, amino acid transport and metabolism. Interestingly some of these results contradict findings in *mecA*$^{+}$ *S. haemolyticus* isolates in public settings as hospital isolates had significantly more genes for defense mechanisms. This could be due to the fact that *S. hominis* isolates were recovered from public settings in West London where *mecA*$^{+}$ *S. haemolyticus* was recovered from public settings in East London. If there were sequences of isolates of *S. hominis* isolates from public areas in East London or *S. haemolyticus* isolates from public areas in West London, then it is possible to determine whether one area in the community would have a higher abundance of unique genes for defense mechanisms. Interestingly, one of the genes identified to be ubiquitous in *mecA*$^{+}$ *S. hominis* isolates from public areas in the community was a spore coat protein. This is interesting, as staphylococci are non-spore forming bacteria; therefore, it is expect that this gene was not responsible for spore-coating formation in *S. hominis* (Pruitt et al., 2007). It is possible that they have high homology to *Bacillus* spore coat protein but may have a different function in staphylococci.

6.5 Conclusion

In this chapter, public setting *mecA*$^+$ *S. hominis* isolates from East and West London were genetically compared to *S. hominis* isolates from different sources from the ENA database. From the phylogenetic analysis, the genetic relatedness of *S. hominis* isolates from public settings was determined to that of other isolates from the ENA database which has not previously done before. Additionally, the antibiotic resistance genes, virulent genes, mobile elements and the pangenome were also analysed. From these analyses, there were novel findings. These findings were:

1. *S. hominis* has an array of virulence genes which have not been fully characterised including the capsule genes. Some of the isolates had a different number of genes that encode for the capsule gene.
2. *MecA*$^+$ *S. hominis* recovered from the public settings were related to isolates obtained from the ENA database recovered from mosquitos, healthy human skin, plants and air samples from residential areas. This shows that public setting isolates may have come from different sources and may have been transmitted via mosquitos or through the air.
3. *S. hominis* recovered from clinical samples belong to a separate genetic lineage and have different accessory genome to that of other isolates which shows these isolates are most likely nosocomial and have not spread into different niches.
4. The majority *S. hominis* recovered from livestock were shown to belong to same genetic lineages and have evolved separately to that of *S. hominis* isolates from different sources and have not spread into different niches based on their accessory genome.
5. Similar to *mecA*$^+$ *S. haemolyticus* isolates, the *mecA*$^+$ *S. hominis* recovered from general public settings had more unique genes that had a function for replication,

recombination and repair compared to *mecA*$^{+}$ isolates recovered from hospitals, indicating that they experience more DNA damage in general public settings than in public areas in hospitals.

Chapter 7: Concluding remarks

7.1 Introduction

Antimicrobial resistance (AMR) is a public health concern. Currently, 700,000 people die worldwide from bacterial infections that no longer respond to antibiotics (O'Neill, 2016). By 2050 it is predicted that the death toll will increase to 10 million as during the post-antibiotic era where simple bacterial infections will no longer be treatable (Bragg et al., 2018; O'Neill, 2016). Therefore, studies looking at the abundance of antibiotic resistant bacteria in different environments and ecological niches and studying their genetic features/lineages and the transmission of AMR genes between humans, livestock and the environment are very important to aid in understanding the evolution of multidrug resistant bacteria (MDR) and the factors driving antibiotic resistance and developing disease.

This is particularly important for staphylococci, as these bacteria are capable of colonising different human body sites (e.g. skin and nostrils) and have the potential to cause infections (Becker et al., 2014). Previous studies have shown that MDR staphylococci have been isolated from clinical samples, people from the community and livestock and wildlife animals, and have been linked to causing infectious diseases (Argudín et al., 2015b; Naimi, 2003; Stefani and Varaldo, 2003; Weese, 2010). However, there is a limited number of studies on the abundance of MDR staphylococci in public settings, and little is known about their clonal lineages/genetic variations and the potential risk they pose to public health (Argudín et al., 2015b; Chen et al., 2017; Conceição et al., 2013; Green et al., 2010; Lutz et al., 2014, 2014; Seng et al., 2017b; Xu et al., 2015). Therefore, in this study, it was set out to determine the levels of MDR staphylococci recovered from high frequency touched surfaces in public settings in East and West London and from public areas in the community and public areas in hospitals. This study

aimed to identify the MDR *mecA*$^{+}$ isolates and determine their genetic lineages and variations they evolved by using whole genome sequencing (WGS) analysis. Using the "One health" approach, WGS data in this study was compared to known isolates obtained from the ENA database that were previously recovered and examined by others, including those from clinical samples, agriculture and companion animals, plants and other environmental isolates to determine if isolates from public settings are a public health risk.

7.2 Findings that link MDR staphylococci recovered from public settings as a potential public health risk and implications

- A total of 600 samples were collected in this study. A large portion of them (46.8%) were MDR staphylococci. These isolates belonged to 11 different species and were resistant to a broad range of antibiotics, which coincides with previous studies (Seng et al., 2017b; Xu et al., 2015). These results show that high frequency touched surfaces in public settings are reservoirs for MDR staphylococci, which have the potential to spread to people from different communities.
- The most commonly found antibiotic staphylococcal isolates were resistant to was penicillin (80.42%), fusidic acid (72.4%), erythromycin (54.5%), amoxicillin (27.8%); tetracycline (26.3%); oxacillin (24.9%); cefoxitin (22.4%); mupirocin (14.6%); gentamycin (9.6%); cefepime (7.1%) and chloramphenicol (5.0%). These results show that if these isolates cause an infection, they may be unable to be treated due to their resistance to commonly used antibiotics.
- There was a significantly higher proportion (P=0.0458) of MDR resistance staphylococci from public areas in hospitals (49.5%) than in general public settings (40.7%) due to there being a higher usage of antibiotics in these areas

(Cantón and Morosini, 2011). This indicates that isolates recovered from hospital settings would be harder to treat with antibiotics compared to isolates from public settings.

- There was a significantly higher proportion (P=0.0002) of MDR resistant staphylococci from public areas in East London (56.7%) compared to West London (50.0%). This may be linked to East London having a higher population density ($9.7x10^3$ km^2; 2017 estimate) compared to West London ($8.^9x10^3$ km^2; 2017 estimate) (Park, 2017; Bruinsma et al., 2003).This shows that people who live in areas of high population density are at higher risk to be infected by MDR staphylococci.

  MecA$^+$ *S. epidermidis*, *S. haemolyticus* and *S. hominis* from public settings had intact phage sequences in their genome that is known to infect other *Staphylococcus* species from other species including *S. aureus*. This suggest it is possible for CoNS in public environment to horizontally receive virulence genes (staphylokinases, superantigens, Panton-Valentine leucocidin) in *S. aureus* that is found to be transferred to other *S. aureus* isolates via phages which could make these CoNS isolates more virulent and a significant risk to public health (Goerke et al., 2009).

- Some of the *S. epidermidis* and *S. haemolyticus mecA*$^+$ isolates in this study belonged to the same genetic lineages as the isolates obtained from the ENA database that were recovered from clinical blood, eye and urine samples which have not been previously reported. Additionally, *mecA*$^+$ *S. epidermidis* and *S. haemolyticus* shared similar virulent gene profiles and other accessory genes to that of known clinical isolates from the ENA database. These findings show that isolates recovered from high-frequency touched surfaces have the potential to

cause infection, which could be untreatable due to being MDR and hence pose a public health risk.

- *MecA*$^{+}$ *S. epidermidis*, S. *haemolyticus* and *S. hominis* isolates from public settings shared the same or similar antibiotic resistance and virulence gene profiles to that of isolates from the ENA database recovered from clinical samples. This implies that the *mecA*$^{+}$ isolates from public settings may be able to cause infections in similar manner to that of clinical isolates which may be hard to treat as they have resistance to antibiotics which are mainly used in hospitals.

7.3 Genetic variations identified in staphylococci recovered from public settings

Apart from attempting to understand the levels of MDR staphylococci recovered from high-frequency touched surfaces and the potential health risk they pose. Additionally, genetic variations were identified that have previously not been described.

- In this study, it was identified that the isolates recovered from public settings had a large array of SCC*mec* types many of which had not been reported due to them having additional or novel *ccr* complex elements which have only previously been identified in clinical and community-associated isolates (Chen et al., 2017b). This shows there may be many genetic variations that have been undiscovered in staphylococci as they may only be present in isolates from public settings.
- *MecA*$^{+}$ *S. epidermidis*, *S. haemolyticus* and *S. hominis* isolates in this study were genetically diverse, as shown in phylogenetic and pangenome studies. Apart from some of the *mecA*$^{+}$ *S. epidermidis* and *S. haemolyticus* isolates from public settings being phylogenetically related to clinical isolates the other *mecA*$^{+}$ isolates, they were also genetically related to isolates from the ENA database that were recovered from healthy humans, livestock, plants and other environmental isolates. Additionally, *mecA*$^{+}$ *S. haemolyticus* isolate was genetically related to the ENA

database isolates recovered from companion animals. For *mecA*$^{+}$ S. *hominis* they were genetically not related to isolate from the ENA database recovered from livestock but were genetically similar to isolate's recovered from healthy humans, mosquitos, air samples in residential areas and environmental isolates. This shows that *Staphylococcus* can be spread into public settings either via human contact, food, animal born vectors or through the air.

- *MecA*$^{+}$ *S. epidermidis, S. haemolyticus* or *S. hominis* recovered from public settings in this study were clustered with isolates recovered from many different sources (the accessory genome phylogenetic analyses and pangenome). These results show that isolates recovered from public settings in this study are likely to be related to isolates that were recovered from different environments. This suggests that high-frequency touched surfaces in public settings carry different species of MDR staphylococci belonging to different lineages, potentially capable of transferring genes between different species from different sources.
- Using pangenome analyses, it was identified that there were unique genes in *S. epidermidis*, *S. haemolyticus* and *S. hominis* that were only found in the *mecA*$^{+}$ isolates recovered from the public settings; however, there were no genes identified which would have been detected in all samples or genes responsible for adaptation for survival in these environments.
- Compared to isolates from the ENA database recovered from a different source, isolates recovered from public settings in this study possessed unique gene encoding for replication, recombination and repair. This suggests that staphylococci in public settings may be exposed to different environmental factors (UV light, ionising radiation and genotoxic chemicals), which could cause DNA damage (Žgur-Bertok, 2013).

7.4 Limitations of this study

To avoid sampling bias, up to 10 colonies of different morphology were randomly chosen for each swab taken. This was not always possible as some would have less than 10 on the plate with many of them being similar in their colony morphology; therefore, some colonies picked that were recovered from the same swabs could have had clonal identities. Due to cost and time restrictions, it was limited to only do WGS analyses of *mecA*$^{+}$ isolates. If it was possible to WGS on all the MDR staphylococci, then it would be possible to identify isolates recovered from the same swab were clones as well as better understand the proportion of the isolates that pose a public health risk.

During this study there were limiting factors, due to time and resources, to only looking at the antibiotic resistance phenotype; therefore it was not possible to determine the virulence phenotype of the isolates recovered in this study. Therefore, the virulence gene profiles of the isolates from public settings in comparison with other isolates recovered from different sources are only a prediction of their virulence ability as it is possible that these genes are not expressed (Kwong et al., 2015).

The main limitation of the WGS analyses in this study was the lack of WGS data for coagulase-negative staphylococcal species recovered from other sources available in the ENA database which would allow us to conduct a sophisticated comparison of isolates. There were data available for isolates recovered from clinical samples and animals, but there was a limited number of isolates recovered from plants and other environmental sources. Additionally, for *S. hominis* genome comparison studies, there were only 40 accessible isolates in total that were WGS. This is most likely due to them not being as commonly found as *S. epidermidis* and *S. haemolyticus* to cause infections in humans and animals. For *S. warneri* and *S. cohnii* there were only 3 *mecA*$^{+}$ isolates that were whole

genome sequenced; therefore, it was not possible to have a good comparison with other isolates from different sources

7.5 Future work

This research has built on previous studies of staphylococci from public settings (beaches, buses, hotels, restrooms and built-up areas) by incorporating WGS; but still requires further research to expand understanding of MDR staphylococci from public settings (Conceição et al., 2013; Lutz et al., 2014; Mkrtchyan et al., 2013; Seng et al., 2017b; Xu et al., 2015). For the samples collected, this research is a snapshot of a single day, whereas in future studies, the same area can be sampled multiple times in a day or over different seasons. This will help to determine if weather; changes in the influx of people to these areas and time points before and after surface disinfectant might be driving factors in the proportion of MDR staphylococci found in public setting areas; what species are present; the difference in antibiotic resistance phenotype and genotype profile and if there is a difference in their genetic lineages.

Along with WGS analyses and antibiotic resistance analyses, further studies could look at these isolates virulence ability by performing phenotyping assay or to use RNA sequencing to determine if the predicted virulence genes are expressed. This will help to better understand how serious the risk is these isolates pose to public health.

7.6 Final Statement

This thesis has helped build on the knowledge of previous studies on MDR staphylococci recovered from public settings and incorporated WGS analysis to determine their genetic variations and lineages. The growing threat of AMR bacteria to public health is concern internationally and the research has shown that isolates recovered from public settings are a public health risk which, in the case of developing a disease, could be untreatable

with currently available antibiotics. These high-frequency touched surfaces in public settings are reservoirs for different lineages of staphylococci that were phylogenetically related with isolates recovered from clinical infections cases, healthy humans, livestock and companion animals and in their accessory genome possessed genes that have transferred across from many bacterial isolates originated in different ecological niches.

Reference

Aanensen, D.M., Feil, E.J., Holden, M.T.G., Dordel, J., Yeats, C.A., Fedosejev, A., Goater, R., Castillo-Ramírez, S., Corander, J., Colijn, C., Chlebowicz, M.A., Schouls, L., Heck, M., Pluister, G., Ruimy, R., Kahlmeter, G., Åhman, J., Matuschek, E., Friedrich, A.W., Parkhill, J., Bentley, S.D., Spratt, B.G., Grundmann, H., 2016. Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive Staphylococcus aureus in Europe. mBio 7, e00444-16. https://doi.org/10.1128/mBio.00444-16

Ahmad, N., Ruzan, I.N., Abd Ghani, M.K., Hussin, A., Nawi, S., Aziz, M.N., Maning, N., Eow, V.L.K., 2009. Characteristics of community- and hospital-acquired meticillin-resistant Staphylococcus aureus strains carrying SCCmec type IV isolated in Malaysia. J. Med. Microbiol. 58, 1213–1218. https://doi.org/10.1099/jmm.0.011353-0

Akanbi, O.E., Njom, H.A., Fri, J., Otigbu, A.C., Clarke, A.M., 2017. Antimicrobial Susceptibility of Staphylococcus aureus Isolated from Recreational Waters and Beach Sand in Eastern Cape Province of South Africa. Int J Environ Res Public Health 14. https://doi.org/10.3390/ijerph14091001

Alcalde-Rico, M., Hernando-Amado, S., Blanco, P., Martínez, J.L., 2016. Multidrug Efflux Pumps at the Crossroad between Antibiotic Resistance and Bacterial Virulence. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.01483

Alirol, E., Wi, T.E., Bala, M., Bazzo, M.L., Chen, X.-S., Deal, C., Dillon, J.-A.R., Kularatne, R., Heim, J., Huijsduijnen, R.H. van, Hook, E.W., Lahra, M.M., Lewis, D.A., Ndowa, F., Shafer, W.M., Tayler, L., Workowski, K., Unemo, M., Balasegaram, M., 2017. Multidrug-resistant gonorrhea: A research and

development roadmap to discover new medicines. PLOS Medicine 14, e1002366. https://doi.org/10.1371/journal.pmed.1002366

Aloush, V., Navon-Venezia, S., Seigman-Igra, Y., Cabili, S., Carmeli, Y., 2006. Multidrug-Resistant Pseudomonas aeruginosa: Risk Factors and Clinical Impact. Antimicrobial Agents and Chemotherapy 50, 43–48. https://doi.org/10.1128/AAC.50.1.43-48.2006

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., Pevzner, P.A., 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics 32, 3380–3387. https://doi.org/10.1093/bioinformatics/btw493

Antunes, L.C.S., Poppleton, D.I., Klingl, A., Criscuolo, A., Dupuy, B., Brochier-Armanet, C., Beloin, C., Gribaldo, S., 2016. Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes, in: ELife. https://doi.org/10.7554/eLife.14589

Argemi, X., Nanoukon, C., Affolabi, D., Keller, D., Hansmann, Y., Riegel, P., Baba-Moussa, L., Prévost, G., 2018. Comparative Genomics and Identification of an Enterotoxin-Bearing Pathogenicity Island, SEPI-1/SECI-1, in Staphylococcus epidermidis Pathogenic Strains. Toxins (Basel) 10. https://doi.org/10.3390/toxins10030093

Argimón, S., Abudahab, K., Goater, R.J.E., Fedosejev, A., Bhai, J., Glasner, C., Feil, E.J., Holden, M.T.G., Yeats, C.A., Grundmann, H., Spratt, B.G., Aanensen, D.M., 2016. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genom 2. https://doi.org/10.1099/mgen.0.000093

Argudín, M.A., Vanderhaeghen, W., Butaye, P., 2015. Antimicrobial resistance and population structure of Staphylococcus epidermidis recovered from pig farms in Belgium. Vet. J. 203, 302–308. https://doi.org/10.1016/j.tvjl.2015.01.008

Veterinary Microbiology 178, 105–113. https://doi.org/10.1016/j.vetmic.2015.04.019

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44, W16–W21. https://doi.org/10.1093/nar/gkw387

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., Hiramatsu, K., 2009. Complete Genome Sequence of Macrococcus caseolyticus Strain JSCS5402, Reflecting the Ancestral Genome of the Human-Pathogenic Staphylococci. J Bacteriol 191, 1180–1190. https://doi.org/10.1128/JB.01058-08

Barbier, F., Ruppé, E., Hernandez, D., Lebeaux, D., Francois, P., Felix, B., Desprez, A., Maiga, A., Woerther, P.-L., Gaillard, K., Jeanrot, C., Wolff, M., Schrenzel, J., Andremont, A., Ruimy, R., 2010. Methicillin-resistant coagulase-negative staphylococci in the community: high homology of SCCmec IVa between Staphylococcus epidermidis and major clones of methicillin-resistant Staphylococcus aureus. J. Infect. Dis. 202, 270–281. https://doi.org/10.1086/653483

Barbu, E.M., Ganesh, V.K., Gurusiddappa, S., Mackenzie, R.C., Foster, T.J., Sudhof, T.C., Höök, M., 2010. β-Neurexin Is a Ligand for the Staphylococcus aureus MSCRAMM SdrC. PLOS Pathogens 6, e1000726. https://doi.org/10.1371/journal.ppat.1000726

Barros, E.M., Ceotto, H., Bastos, M.C.F., dos Santos, K.R.N., Giambiagi-deMarval, M., 2012. Staphylococcus haemolyticus as an Important Hospital Pathogen and

Carrier of Methicillin Resistance Genes. J Clin Microbiol 50, 166–168. https://doi.org/10.1128/JCM.05563-11

Basaglia, G., Moras, L., Bearz, A., Scalone, S., Paoli, P.D., 2003. Staphylococcus cohnii septicaemia in a patient with colon cancer. J. Med. Microbiol. 52, 101–102. https://doi.org/10.1099/jmm.0.05002-0

Bautista-Trujillo, G.U., Solorio-Rivera, J.L., Renteria-Solorzano, I., Carranza-German, S.I., Bustos-Martinez, J.A., Arteaga-Garibay, R.I., Baizabal-Aguirre, V.M., Cajero-Juarez, M., Bravo-Patino, A., Valdez-Alarcon, J.J., 2013. Performance of culture media for the isolation and identification of Staphylococcus aureus from bovine mastitis. Journal of Medical Microbiology 62, 369–376. https://doi.org/10.1099/jmm.0.046284-0

Bean, D.C., Wigmore, S.M., Wareham, D.W., 2017. Draft Genome Sequence of a Canine Isolate of Methicillin-Resistant Staphylococcus haemolyticus. Genome Announc 5. https://doi.org/10.1128/genomeA.00146-17

Becker, K., Heilmann, C., Peters, G., 2014. Coagulase-Negative Staphylococci. Clin Microbiol Rev 27, 870–926. https://doi.org/10.1128/CMR.00109-13

Becker, K., Ballhausen, B., Kahl, B.C., Köck, R., 2017. The clinical impact of livestock-associated methicillin-resistant Staphylococcus aureus of the clonal complex 398 for humans. Vet. Microbiol. 200, 33–38. https://doi.org/10.1016/j.vetmic.2015.11.013

Bertelli, C., Brinkman, F.S.L., 2018. Improved genomic island predictions with IslandPath-DIMOB. Bioinformatics 34, 2161–2167. https://doi.org/10.1093/bioinformatics/bty095

Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing Group, Lau, B.Y., Hoad, G.,

Winsor, G.L., Brinkman, F.S.L., 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 45, W30–W35. https://doi.org/10.1093/nar/gkx343

Bhargava, K., Zhang, Y., 2012. Multidrug-resistant coagulase-negative Staphylococci in food animals. Journal of Applied Microbiology 113, 1027–1036. https://doi.org/10.1111/j.1365-2672.2012.05410.x

Bhatta, D.R., Hamal, D., Shrestha, R., Hosuru Subramanya, S., Baral, N., Singh, R.K., Nayak, N., Gokhale, S., 2018. Bacterial contamination of frequently touched objects in a tertiary care hospital of Pokhara, Nepal: how safe are our hands? Antimicrob Resist Infect Control 7. https://doi.org/10.1186/s13756-018-0385-2

Blake, J.E., Metcalfe, M.A., 2001. A Shared Noncapsular Antigen Is Responsible for False-Positive Reactions by Staphylococcus epidermidis in Commercial Agglutination Tests for Staphylococcus aureus. J Clin Microbiol 39, 544–550. https://doi.org/10.1128/JCM.39.2.544-550.2001

Borderon, E., Horodniceanu, T., 1978. Metabolically deficient dwarf-colony mutants of Escherichia coli: deficiency and resistance to antibiotics of strains isolated from urine culture. J Clin Microbiol 8, 629–634.

Bosi, E., Fani, R., Fondi, M., 2015. Defining orthologs and pangenome size metrics. Methods Mol. Biol. 1231, 191–202. https://doi.org/10.1007/978-1-4939-1720-4_13

Bosi, E., Monk, J.M., Aziz, R.K., Fondi, M., Nizet, V., Palsson, B.Ø., 2016. Comparative genome-scale modelling of Staphylococcus aureus strains identifies strain-specific metabolic capabilities linked to pathogenicity. PNAS 113, E3801–E3809. https://doi.org/10.1073/pnas.1523199113

Boundy, S., Safo, M.K., Wang, L., Musayev, F.N., O'Farrell, H.C., Rife, J.P., Archer, G.L., 2013. Characterization of the Staphylococcus aureus rRNA methyltransferase encoded by orfX, the gene containing the staphylococcal chromosome Cassette mec (SCCmec) insertion site. J. Biol. Chem. 288, 132–140. https://doi.org/10.1074/jbc.M112.385138

Bragg, R.R., Meyburgh, C.M., Lee, J.-Y., Coetzee, M., 2018. Potential Treatment Options in a Post-antibiotic Era. Adv. Exp. Med. Biol. 1052, 51–61. https://doi.org/10.1007/978-981-10-7572-8_5

Brooks, B., Olm, M.R., Firek, B.A., Baker, R., Thomas, B.C., Morowitz, M.J., Banfield, J.F., 2017. Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome. Nat Commun 8, 1–7. https://doi.org/10.1038/s41467-017-02018-w

Bruinsma, N., Hutchinson, J.M., van den Bogaard, A.E., Giamarellou, H., Degener, J., Stobberingh, E.E., 2003. Influence of population density on antibiotic resistance. J. Antimicrob. Chemother. 51, 385–390.

Bryant, J., Chewapreecha, C., Bentley, S.D., 2012. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. Future microbiology 7, 1283. https://doi.org/10.2217/fmb.12.108

Buelow, E., Bayjanov, J.R., Majoor, E., Willems, R.J., Bonten, M.J., Schmitt, H., van Schaik, W., 2018. Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. FEMS Microbiol Ecol 94. https://doi.org/10.1093/femsec/fiy087

Burts, M.L., DeDent, A.C., Missiakas, D.M., 2008. EsaC substrate for the ESAT-6 Secretion Pathway and its role in persistent infections of S. aureus. Mol Microbiol 69, 736–746. https://doi.org/10.1111/j.1365-2958.2008.06324.x

Burts, M.L., Williams, W.A., DeBord, K., Missiakas, D.M., 2005. EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of Staphylococcus aureus infections. Proc Natl Acad Sci U S A 102, 1169–1174. https://doi.org/10.1073/pnas.0405620102

Büttner, H., Mack, D., Rohde, H., 2015. Structural basis of Staphylococcus epidermidis biofilm formation: mechanisms and molecular interactions. Front Cell Infect Microbiol 5. https://doi.org/10.3389/fcimb.2015.00014

Caboche, S., Audebert, C., Lemoine, Y., Hot, D., 2014. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. BMC Genomics 15, 264. https://doi.org/10.1186/1471-2164-15-264

Cadieux, B., Vijayakumaran, V., Bernards, M.A., McGavin, M.J., Heinrichs, D.E., 2014. Role of Lipase from Community-Associated Methicillin-Resistant Staphylococcus aureus Strain USA300 in Hydrolyzing Triglycerides into Growth-Inhibitory Free Fatty Acids. Journal of Bacteriology 196, 4044–4056. https://doi.org/10.1128/JB.02044-14

Calkins, S., Couger, M.B., Jackson, C., Zandler, J., Hudgins, G.C., Hanafy, R.A., Budd, C., French, D.P., Hoff, W.D., Youssef, N., 2016. Draft genome sequence of Staphylococcus hominis strain Hudgins isolated from human skin implicates metabolic versatility and several virulence determinants. Genom Data 10, 91–96. https://doi.org/10.1016/j.gdata.2016.10.003

Campbell, I., 2007. Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. Statistics in Medicine 26, 3661–3675. https://doi.org/10.1002/sim.2832

to antibiotics. FEMS Microbiol. Rev. 35, 977–991. https://doi.org/10.1111/j.1574-6976.2011.00295.x

Cao, Z., Casabona, M.G., Kneuper, H., Chalmers, J.D., Palmer, T., 2016. The type VII secretion system of Staphylococcus aureus secretes a nuclease toxin that targets competitor bacteria. Nat Microbiol 2, 16183. https://doi.org/10.1038/nmicrobiol.2016.183

Caputo, A., Fournier, P.-E., Raoult, D., 2019. Genome and pan-genome analysis to classify emerging bacteria. Biology Direct 14, 5. https://doi.org/10.1186/s13062-019-0234-0

Cavanagh, J. P., Hjerde, E., Holden, M.T., Kahlke, T., Klingenberg, C., Flægstad, T., Parkhill, J., Bentley, S.D., Sollid, J.U., 2014. Whole-genome sequencing reveals clonal expansion of multiresistant Staphylococcus haemolyticus in European hospitals. J Antimicrob Chemother 69, 2920–2927. https://doi.org/10.1093/jac/dku271

Chambers, H.F., DeLeo, F.R., 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. Nature Reviews Microbiology 7, 629–641. https://doi.org/10.1038/nrmicro2200

Chan, K.G., Ng, K.T., Chong, T.M., Pang, Y.K., Kamarulzaman, A., Yin, W.F., Tee, K.K., 2015. Antibiotic Resistant and Virulence Determinants of Staphylococcus haemolyticus C10A as Revealed by Whole Genome Sequencing. J Genomics 3, 72–74. https://doi.org/10.7150/jgen.12574

Chaudhry, V., Patil, P.B., 2016. Genomic investigation reveals evolution and lifestyle adaptation of endophytic *Staphylococcus epidermidis*. Scientific Reports 6, 19263. https://doi.org/10.1038/srep19263

Chaves, F., Garcia-Alvarez, M., Sanz, F., Alba, C., Otero, J.R., 2005. Nosocomial Spread of a Staphylococcus hominis subsp. novobiosepticus Strain Causing Sepsis in a

Neonatal Intensive Care Unit. Journal of Clinical Microbiology 43, 4877–4879. https://doi.org/10.1128/JCM.43.9.4877-4879.2005

Checinska Sielaff, A., Singh, N.K., Allen, J.E., Thissen, J., Jaing, C., Venkateswaran, K., 2016. Draft Genome Sequences of Biosafety Level 2 Opportunistic Pathogens Isolated from the Environmental Surfaces of the International Space Station. Genome Announc 4. https://doi.org/10.1128/genomeA.01263-16

Chen, H.-J., Chang, Y.-C., Tsai, J.-C., Hung, W.-C., Lin, Y.-T., You, S.-J., Tseng, S.-P., Teng, L.-J., 2013. New Structure of Phage-Related Islands Carrying fusB and a Virulence Gene in Fusidic Acid-Resistant Staphylococcus epidermidis. Antimicrob Agents Chemother 57, 5737–5739. https://doi.org/10.1128/AAC.01433-13

Chen, X.-P., Li, W.-G., Zheng, H., Du, H.-Y., Zhang, Li, Zhang, Lei, Che, J., Wu, Y., Liu, S.-M., Lu, J.-X., 2017. Extreme diversity and multiple SCCmec elements in coagulase-negative Staphylococcus found in the Clinic and Community in Beijing, China. Annals of Clinical Microbiology and Antimicrobials 16. https://doi.org/10.1186/s12941-017-0231-z

Cheng, V.W.T., Zhang, G., Oyedotun, K.S., Ridgway, D., Ellison, M.J., Weiner, J.H., 2013. Complete Genome of the Solvent-Tolerant Staphylococcus warneri Strain SG1. Genome Announc 1. https://doi.org/10.1128/genomeA.00038-13

Cho, S.-H., Naber, K., Hacker, J., Ziebuhr, W., 2002. Detection of the icaADBC gene cluster and biofilm formation in Staphylococcus epidermidis isolates from catheter-related urinary tract infections. Int. J. Antimicrob. Agents 19, 570–575.

Christensen, G.D., Simpson, W.A., Younger, J.J., Baddour, L.M., Barrett, F.F., Melton, D.M., Beachey, E.H., 1985. Adherence of coagulase-negative staphylococci to

plastic tissue culture plates: a quantitative model for the adherence of staphylococci to medical devices. J Clin Microbiol 22, 996–1006.

Chu, V.H., Woods, C.W., Miro, J.M., Hoen, B., Cabell, C.H., Pappas, P.A., Federspiel, J., Athan, E., Stryjewski, M.E., Nacinovich, F., Marco, F., Levine, D.P., Elliott, T.S., Fortes, C.Q., Tornos, P., Gordon, D.L., Utili, R., Delahaye, F., Corey, G.R., Fowler, V.G., 2008. Emergence of Coagulase-Negative Staphylococci as a Cause of Native Valve Endocarditis. Clin Infect Dis 46, 232–242. https://doi.org/10.1086/524666

Clark, A.E., Kaleta, E.J., Arora, A., Wolk, D.M., 2013. Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. Clin. Microbiol. Rev. 26, 547–603. https://doi.org/10.1128/CMR.00072-12).

Coburn, P.S., Pillar, C.M., Jett, B.D., Haas, W., Gilmore, M.S., 2004. Enterococcus faecalis senses target cells and in response expresses cytolysin. Science 306, 2270–2272. https://doi.org/10.1126/science.1103996

Cole, K., Foster, D., Russell, J.E., Golubchik, T., Llewelyn, M., Wilson, D.J., Crook, D., Paul, J., 2019a. Draft Genome Sequences of 64 Type Strains of 50 Species and 25 Subspecies of the Genus Staphylococcus Rosenbach 1884. Microbiol Resour Announc 8, e00062-19. https://doi.org/10.1128/MRA.00062-19

Conceição, T., Diamantino, F., Coelho, C., de Lencastre, H., Aires-de-Sousa, M., 2013. Contamination of Public Buses with MRSA in Lisbon, Portugal: A Possible Transmission Route of Major MRSA Clones within the Community. PLoS One 8. https://doi.org/10.1371/journal.pone.0077812

Conlan, S., Mijares, L.A., Becker, J., Blakesley, R.W., Bouffard, G.G., Brooks, S., Coleman, H., Gupta, J., Gurson, N., Park, M., Schmidt, B., Thomas, P.J., Otto,

M., Kong, H.H., Murray, P.R., Segre, J.A., 2012. Staphylococcus epidermidis pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. Genome Biol 13, R64. https://doi.org/10.1186/gb-2012-13-7-r64

Costa, S.S., Sobkowiak, B., Parreira, R., Edgeworth, J.D., Viveiros, M., Clark, T.G., Couto, I., 2019. Genetic Diversity of norA, Coding for a Main Efflux Pump of Staphylococcus aureus. Front. Genet. 9. https://doi.org/10.3389/fgene.2018.00710

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., Knight, R., 2009. Bacterial Community Variation in Human Body Habitats Across Space and Time. Science 326, 1694–1697. https://doi.org/10.1126/science.1177486

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

Croxatto, A., Prod'hom, G., Greub, G., 2012. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. FEMS Microbiol Rev 36, 380–407. https://doi.org/10.1111/j.1574-6976.2011.00298.x

Cunha, B.A., Esrick, M.D., LaRusso, M., 2007. Staphylococcus hominis native mitral valve bacterial endocarditis (SBE) in a patient with hypertrophic obstructive cardiomyopathy. Heart & Lung 36, 380–382. https://doi.org/10.1016/j.hrtlng.2006.11.002

Cunnion, K.M., Zhang, H.-M., Frank, M.M., 2003. Availability of complement bound to Staphylococcus aureus to interact with membrane complement receptors influences efficiency of phagocytosis. Infect. Immun. 71, 656–662.

Czekaj, T., Ciszewski, M., Szewczyk, E.M., 2015. Staphylococcus haemolyticus - an emerging threat in the twilight of the antibiotics age. Microbiology (Reading, Engl.) 161, 2061–2068. https://doi.org/10.1099/mic.0.000178

Daniel, A., Bonnen, P.E., Fischetti, V.A., 2007. First Complete Genome Sequence of Two Staphylococcus epidermidis Bacteriophages. J Bacteriol 189, 2086–2100. https://doi.org/10.1128/JB.01637-06

Dantas, G., Sommer, M., 2014. How to Fight Back Against Antibiotic Resistance. Am. Sci. 102, 42. https://doi.org/10.1511/2014.106.42

David, M.Z., Daum, R.S., 2010. Community-Associated Methicillin-Resistant Staphylococcus aureus: Epidemiology and Clinical Consequences of an Emerging Epidemic. Clinical Microbiology Reviews 23, 616–687. https://doi.org/10.1128/CMR.00081-09

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., Wright, G.D., 2011. Antibiotic resistance is ancient. Nature 477, 457–461. https://doi.org/10.1038/nature10388

Degener, J.E., Heck, M.E., van Leeuwen, W.J., Heemskerk, C., Crielaard, A., Joosten, P., Caesar, P., 1994. Nosocomial infection by Staphylococcus haemolyticus and typing methods for epidemiological study. J Clin Microbiol 32, 2260–2265.

Deghorain, M., Bobay, L.-M., Smeesters, P.R., Bousbata, S., Vermeersch, M., Perez-Morga, D., Drèze, P.-A., Rocha, E.P.C., Touchon, M., Melderen, L.V., 2012. Characterization of Novel Phages Isolated in Coagulase-Negative Staphylococci Reveals Evolutionary Relationships with Staphylococcus aureus Phages. Journal of Bacteriology 194, 5829–5839. https://doi.org/10.1128/JB.01085-12

Deghorain, M., Van Melderen, L., 2012. The Staphylococci Phages Family: An Overview. Viruses 4, 3316–3335. https://doi.org/10.3390/v4123316

Deplano, A., Vandendriessche, S., Nonhoff, C., Dodémont, M., Roisin, S., Denis, O., 2016. National surveillance of Staphylococcus epidermidis recovered from bloodstream infections in Belgian hospitals. J. Antimicrob. Chemother. 71, 1815–1819. https://doi.org/10.1093/jac/dkw086

Diep, B.A., Gill, S.R., Chang, R.F., Phan, T.H., Chen, J.H., Davidson, M.G., Lin, F., Lin, J., Carleton, H.A., Mongodin, E.F., Sensabaugh, G.F., Perdreau-Remington, F., 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. Lancet 367, 731–739. https://doi.org/10.1016/S0140-6736(06)68231-7

Dolk, F.C.K., Pouwels, K.B., Smith, D.R.M., Robotham, J.V., Smieszek, T., 2018. Antibiotics in primary care in England: which antibiotics are prescribed and for which conditions? J Antimicrob Chemother 73, ii2–ii10. https://doi.org/10.1093/jac/dkx504

Dorado-García, A., Bos, M.E., Graveland, H., Cleef, B.A.V., Verstappen, K.M., Kluytmans, J.A., Wagenaar, J.A., Heederik, D.J., 2013. Risk factors for persistence of livestock-associated MRSA and environmental exposure in veal calf farmers and their family members: an observational longitudinal study. BMJ Open 3. https://doi.org/10.1136/bmjopen-2013-003272

dos Santos, D.C.M., da Costa, T.M., Rabello, R.F., Alves, F.A., de Mondino, S.S.B., 2015. Mannitol-negative methicillin-resistant Staphylococcus aureus from nasal swab specimens in Brazil. Braz J Microbiol 46, 531–533. https://doi.org/10.1590/S1517-838246220140179

Elements (IWG-SCC), I.W.G. on the C. of S.C.C., 2009. Classification of Staphylococcal Cassette Chromosome mec (SCCmec): Guidelines for Reporting Novel SCCmec Elements. Antimicrobial Agents and Chemotherapy 53, 4961–4967. https://doi.org/10.1128/AAC.00579-09

Enault, F., Briet, A., Bouteille, L., Roux, S., Sullivan, M.B., Petit, M.-A., 2017. Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. ISME J 11, 237–247. https://doi.org/10.1038/ismej.2016.90

Fajardo, A., Martínez-Martín, N., Mercadillo, M., Galán, J.C., Ghysels, B., Matthijs, S., Cornelis, P., Wiehlmann, L., Tümmler, B., Baquero, F., Martínez, J.L., 2008. The Neglected Intrinsic Resistome of Bacterial Pathogens. PLoS ONE 3, e1619. https://doi.org/10.1371/journal.pone.0001619

Fernandes Dos Santos, F., Mendonça, L.C., Reis, D.R. de L., Guimarães, A. de S., Lange, C.C., Ribeiro, J.B., Machado, M.A., Brito, M.A.V.P., 2016. Presence of mecA-positive multidrug-resistant Staphylococcus epidermidis in bovine milk samples in Brazil. J. Dairy Sci. 99, 1374–1382. https://doi.org/10.3168/jds.2015-9931

Feßler, A.T., Billerbeck, C., Kadlec, K., Schwarz, S., 2010. Identification and characterization of methicillin-resistant coagulase-negative staphylococci from bovine mastitis. J Antimicrob Chemother 65, 1576–1582. https://doi.org/10.1093/jac/dkq172

Flahaut, S., Vinogradov, E., Kelley, K.A., Brennan, S., Hiramatsu, K., Lee, J.C., 2008. Structural and Biological Characterization of a Capsular Polysaccharide Produced by Staphylococcus haemolyticus. Journal of Bacteriology 190, 1649–1657. https://doi.org/10.1128/JB.01648-07

Foster, T., 1996. Staphylococcus, in: Baron, S. (Ed.), Medical Microbiology. University of Texas Medical Branch at Galveston, Galveston (TX).

Fredheim, E.G.A., Klingenberg, C., Rohde, H., Frankenberger, S., Gaustad, P., Flægstad, T., Sollid, J.E., 2009. Biofilm Formation by Staphylococcus haemolyticus. Journal of Clinical Microbiology 47, 1172–1180. https://doi.org/10.1128/JCM.01891-08

Fritze, D., 2004. Taxonomy of the genus bacillus and related genera: the aerobic endospore-forming bacteria. Phytopathology 94, 1245–1248. https://doi.org/10.1094/PHYTO.2004.94.11.1245

Froggatt, J.W., Johnston, J.L., Galetto, D.W., Archer, G.L., 1989. Antimicrobial resistance in nosocomial isolates of Staphylococcus haemolyticus. Antimicrob Agents Chemother 33, 460–466.

Galac, M.R., Hoang, T.-M., Ong, A.C., Hinkle, M., O'Gara, J.P., Fey, P.D., 2019. Complete Genome Sequence of Staphylococcus epidermidis CSF41498. Microbiol Resour Announc 8. https://doi.org/10.1128/MRA.01138-18

Gan, H. Y., Gan, H.M., Savka, M.A., Triassi, A.J., Wheatley, M.S., Smart, L.B., Fabio, E.S., Hudson, A.O., 2014. Whole-Genome Sequences of 13 Endophytic Bacteria Isolated from Shrub Willow (Salix) Grown in Geneva, New York. Genome Announcements 2. https://doi.org/10.1128/genomeA.00288-14

Gao, M., Glenn, A.E., Blacutt, A.A., Gold, S.E., 2017. Fungal Lactamases: Their Occurrence and Function. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.01775

Garneau-Tsodikova, S., Labby, K.J., 2016. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. Medchemcomm 7, 11–27. https://doi.org/10.1039/C5MD00344J

Gaupp, R., Ledala, N., Somerville, G.A., 2012. Staphylococcal response to oxidative stress. Front Cell Infect Microbiol 2. https://doi.org/10.3389/fcimb.2012.00033

Gazzola, S., Pietta, E., Bassi, D., Fontana, C., Puglisi, E., Cappa, F., Cocconcelli, P.S., 2013. Draft Genome Sequence of Vancomycin-Heteroresistant Staphylococcus epidermidis Strain UC7032, Isolated from Food. Genome Announc 1. https://doi.org/10.1128/genomeA.00709-13

Ghasemian, E., Naghoni, A., Tabaraie, B., Tabaraie, T., 2012. In vitro susceptibility of filamentous fungi to copper nanoparticles assessed by rapid XTT colorimetry and agar dilution method. J Mycol Med 22, 322–328. https://doi.org/10.1016/j.mycmed.2012.09.006

Ghaznavi-Rad, E., Shamsudin, M.N., Sekawi, Z., Khoon, L.Y., Aziz, M.N., Hamat, R.A., Othman, N., Chong, P.P., Belkum, A. van, Ghasemzadeh-Moghaddam, H., Neela, V., 2010. Predominance and Emergence of Clones of Hospital-Acquired Methicillin-Resistant Staphylococcus aureus in Malaysia. Journal of Clinical Microbiology 48, 867–872. https://doi.org/10.1128/JCM.01112-09

Gibson, J.S., Cobbold, R.N., Trott, D.J., 2010. Characterization of multidrug-resistant Escherichia coli isolated from extraintestinal clinical infections in animals. Journal of Medical Microbiology 59, 592–598. https://doi.org/10.1099/jmm.0.018002-0

Gill, S. R., Fouts, D.E., Archer, G.L., Mongodin, E.F., DeBoy, R.T., Ravel, J., Paulsen, I.T., Kolonay, J.F., Brinkac, L., Beanan, M., Dodson, R.J., Daugherty, S.C., Madupu, R., Angiuoli, S.V., Durkin, A.S., Haft, D.H., Vamathevan, J., Khouri, H., Utterback, T., Lee, C., Dimitrov, G., Jiang, L., Qin, H., Weidman, J., Tran, K., Kang, K., Hance, I.R., Nelson, K.E., Fraser, C.M., 2005. Insights on Evolution of Virulence and Resistance from the Complete Genome Analysis of an Early Methicillin-Resistant Staphylococcus aureus Strain and a Biofilm-Producing

Methicillin-Resistant Staphylococcus epidermidis Strain. Journal of Bacteriology 187, 2426–2438. https://doi.org/10.1128/JB.187.7.2426-2438.2005

Goerke, C., Pantucek, R., Holtfreter, S., Schulte, B., Zink, M., Grumann, D., Bröker, B.M., Doskar, J., Wolz, C., 2009. Diversity of Prophages in Dominant Staphylococcus aureus Clonal Lineages. Journal of Bacteriology 191, 3462–3468. https://doi.org/10.1128/JB.01804-08

Gómez-Sanz, E., Ceballos, S., Ruiz-Ripa, L., Zarazaga, M., Torres, C., 2019. Clonally Diverse Methicillin and Multidrug Resistant Coagulase Negative Staphylococci Are Ubiquitous and Pose Transfer Ability Between Pets and Their Owners. Front. Microbiol. 10. https://doi.org/10.3389/fmicb.2019.00485

González-Zorn, B., Courvalin, P., 2003. vanA-mediated high level glycopeptide resistance in MRSA. The Lancet Infectious Diseases 3, 67–68. https://doi.org/10.1016/S1473-3099(03)00510-3

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. Nat. Rev. Genet. 17, 333–351. https://doi.org/10.1038/nrg.2016.49

Gorwitz, R.J., Kruszon-Moran, D., McAllister, S.K., McQuillan, G., McDougal, L.K., Fosheim, G.E., Jensen, B.J., Killgore, G., Tenover, F.C., Kuehnert, M.J., 2008. Changes in the prevalence of nasal colonization with Staphylococcus aureus in the United States, 2001-2004. J. Infect. Dis. 197, 1226–1234. https://doi.org/10.1086/533494

Graham, J.E., Moore, Jonathan E., Jiru, X., Moore, John E., Goodall, E.A., Dooley, J.S.G., Hayes, V.E.A., Dartt, D.A., Downes, C.S., Moore, T.C.B., 2007. Ocular pathogen or commensal: a PCR-based study of surface bacterial flora in normal and dry

eyes. Invest. Ophthalmol. Vis. Sci. 48, 5616–5623. https://doi.org/10.1167/iovs.07-0588

Green, S.M., Marsh, P., Ahmad, N., Jefferies, J.M.C., Clarke, S.C., 2010. Characterization of community and hospital Staphylococcus aureus isolates in Southampton, UK. Journal of Medical Microbiology 59, 1084–1088. https://doi.org/10.1099/jmm.0.018986-0

Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., Program, N.C.S., Bouffard, G.G., Blakesley, R.W., Murray, P.R., Green, E.D., Turner, M.L., Segre, J.A., 2009. Topographical and Temporal Diversity of the Human Skin Microbiome. Science 324, 1190–1192. https://doi.org/10.1126/science.1171700

Gründling, A., Schneewind, O., 2007. Genes Required for Glycolipid Synthesis and Lipoteichoic Acid Anchoring in Staphylococcus aureus. Journal of Bacteriology 189, 2521–2530. https://doi.org/10.1128/JB.01683-06

Gutiérrez, D., Martínez, B., Rodríguez, A., García, P., 2012. Genomic characterization of two Staphylococcus epidermidis bacteriophages with anti-biofilm potential. BMC Genomics 13, 228. https://doi.org/10.1186/1471-2164-13-228

Hanaki, H., Labischinski, H., Inaba, Y., Kondo, N., Murakami, H., Hiramatsu, K., 1998. Increase in glutamine-non-amidated muropeptides in the peptidoglycan of vancomycin-resistant Staphylococcus aureus strain Mu50. J. Antimicrob. Chemother. 42, 315–320.

Hau, S.J., Haan, J.S., Davies, P.R., Frana, T., Nicholson, T.L., 2018. Antimicrobial Resistance Distribution Differs Among Methicillin Resistant Staphylococcus aureus Sequence Type (ST) 5 Isolates From Health Care and Agricultural Sources. Front Microbiol 9. https://doi.org/10.3389/fmicb.2018.02102

Hernando-Amado, S., Blanco, P., Alcalde-Rico, M., Corona, F., Reales-Calderón, J.A., Sánchez, M.B., Martínez, J.L., 2016. Multidrug efflux pumps as main players in intrinsic and acquired resistance to antimicrobials. Drug Resistance Updates 28, 13–27. https://doi.org/10.1016/j.drup.2016.06.007

Heuck, A.P., Moe, P.C., Johnson, B.B., 2010. The cholesterol-dependent cytolysin family of gram-positive bacterial toxins. Subcell. Biochem. 51, 551–577. https://doi.org/10.1007/978-90-481-8622-8_20

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., Baba, T., 2014. Multi-drug-resistant Staphylococcus aureus and future chemotherapy. J. Infect. Chemother. 20, 593–601. https://doi.org/10.1016/j.jiac.2014.08.001

Hong, J., Kim, J., Kim, B.Y., Park, J.W., Ryu, J.G., Roh, E., 2016. Complete Genome Sequence of Biofilm-Forming Strain Staphylococcus haemolyticus S167. Genome Announc 4. https://doi.org/10.1128/genomeA.00567-16

Hsiao, C.-H., Ong, S.J., Chuang, C.-C., Ma, D.H.K., Huang, Y.-C., 2015. A Comparison of Clinical Features between Community-Associated and Healthcare-Associated Methicillin-Resistant Staphylococcus aureus Keratitis [WWW Document]. Journal of Ophthalmology. https://doi.org/10.1155/2015/923941

Huang, H., Flynn, N.M., King, J.H., Monchaud, C., Morita, M., Cohen, S.H., 2006. Comparisons of Community-Associated Methicillin-Resistant Staphylococcus aureus (MRSA) and Hospital-Associated MSRA Infections in Sacramento, California. J Clin Microbiol 44, 2423–2427. https://doi.org/10.1128/JCM.00254-06

Huebner, J., Goldmann, D.A., 1999. Coagulase-negative staphylococci: role as pathogens. Annu. Rev. Med. 50, 223–236. https://doi.org/10.1146/annurev.med.50.1.223

Hughes, Grant L., Raygoza Garay, J.A., Koundal, V., Rasgon, J.L., Mwangi, M.M., 2016. Genome Sequences of Staphylococcus hominis Strains ShAs1, ShAs2, and ShAs3, Isolated from the Asian Malaria Mosquito Anopheles stephensi. Genome Announc 4. https://doi.org/10.1128/genomeA.00085-16

Hung, W.-C., Chen, H.-J., Lin, Y.-T., Tsai, J.-C., Chen, C.-W., Lu, H.-H., Tseng, S.-P., Jheng, Y.-Y., Leong, K.H., Teng, L.-J., 2015. Skin Commensal Staphylococci May Act as Reservoir for Fusidic Acid Resistance Genes. PLOS ONE 10, e0143106. https://doi.org/10.1371/journal.pone.0143106

Huntington, M.K., Allison, J., Nair, D., 2016. Emerging Vector-Borne Diseases. Am Fam Physician 94, 551–557.

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

Ito, T., Kuwahara, K., Hiramatsu, K., 2007. Staphylococcal cassette chromosome mec(SCC mec) analysis of MRSA. Methods Mol. Biol. 391, 87–102. https://doi.org/10.1007/978-1-59745-468-1_7

Iyer, M.N., Wirostko, W.J., Kim, S.H., Simons, K.B., 2005. Staphylococcus hominis Endophthalmitis Associated With a Capsular Hypopyon. American Journal of Ophthalmology 139, 930–932. https://doi.org/10.1016/j.ajo.2004.10.060

Janda, J.M., Abbott, S.L., 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. Journal of Clinical Microbiology 45, 2761–2764. https://doi.org/10.1128/JCM.01228-07

Jarraud, S., Cozon, G., Vandenesch, F., Bes, M., Etienne, J., Lina, G., 1999. Involvement of Enterotoxins G and I in Staphylococcal Toxic Shock Syndrome and Staphylococcal Scarlet Fever. J Clin Microbiol 37, 2446–2449.

Jarraud, S., Mougel, C., Thioulouse, J., Lina, G., Meugnier, H., Forey, F., Nesme, X., Etienne, J., Vandenesch, F., 2002. Relationships between Staphylococcus aureus Genetic Background, Virulence Factors, agr Groups (Alleles), and Human Disease. Infect Immun 70, 631–641. https://doi.org/10.1128/IAI.70.2.631-641.2002

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R., Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S.L., Wright, G.D., McArthur, A.G., 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 45, D566–D573. https://doi.org/10.1093/nar/gkw1004

Jiang, Saiping, Zheng, B., Ding, W., Lv, L., Ji, J., Zhang, H., Xiao, Y., Li, L., 2012. Whole-Genome Sequence of Staphylococcus hominis, an Opportunistic Pathogen. Journal of Bacteriology 194, 4761–4762. https://doi.org/10.1128/JB.00991-12

Kashuba, E., Dmitriev, A.A., Kamal, S.M., Melefors, O., Griva, G., Römling, U., Ernberg, I., Kashuba, V., Brouchkov, A., 2017. Ancient permafrost staphylococci carry antibiotic resistance genes. Microb Ecol Health Dis 28. https://doi.org/10.1080/16512235.2017.1345574

Kaya, H., Hasman, H., Larsen, J., Stegger, M., Johannesen, T.B., Allesøe, R.L., Lemvigh, C.K., Aarestrup, F.M., Lund, O., Larsen, A.R., 2018. SCCmecFinder, a Web-Based Tool for Typing of Staphylococcal Cassette Chromosome mec in

Staphylococcus aureus Using Whole-Genome Sequence Data. mSphere 3. https://doi.org/10.1128/mSphere.00612-17

Kaye, K.S., Gold, H.S., Schwaber, M.J., Venkataraman, L., Qi, Y., De Girolami, P.C., Samore, M.H., Anderson, G., Rasheed, J.K., Tenover, F.C., 2004. Variety of beta-lactamases produced by amoxicillin-clavulanate-resistant Escherichia coli isolated in the northeastern United States. Antimicrob. Agents Chemother. 48, 1520–1525.

Kehrenberg, C., Ojo, K.K., Schwarz, S., 2004. Nucleotide sequence and organization of the multiresistance plasmid pSCFS1 from Staphylococcus sciuri. J. Antimicrob. Chemother. 54, 936–939. https://doi.org/10.1093/jac/dkh457

Kerschner, H., Harrison, E.M., Hartl, R., Holmes, M.A., Apfalter, P., 2014. First report of mecC MRSA in human samples from Austria: molecular characteristics and clinical data. New Microbes New Infect 3, 4–9. https://doi.org/10.1016/j.nmni.2014.11.001

khelissa, S.O., Jama, C., Abdallah, M., Boukherroub, R., Faille, C., Chihib, N.-E., 2017. Effect of incubation duration, growth temperature, and abiotic surface type on cell surface properties, adhesion and pathogenicity of biofilm-detached Staphylococcus aureus cells. AMB Express 7. https://doi.org/10.1186/s13568-017-0492-0

Kiedrowski, M.R., Crosby, H.A., Hernandez, F.J., Malone, C.L., Ii, J.O.M., Horswill, A.R., 2014. Staphylococcus aureus Nuc2 Is a Functional, Surface-Attached Extracellular Nuclease. PLOS ONE 9, e95574. https://doi.org/10.1371/journal.pone.0095574

Kim, B., Kim, J., Park, H., Park, J., 2016. Draft Genome Sequence of Toluene-Resistant Staphylococcus epidermidis SNUT. Genome Announc 4. https://doi.org/10.1128/genomeA.00057-16

Klibi, A., Maaroufi, A., Torres, C., Jouini, A., 2018. Detection and characterization of methicillin-resistant and susceptible coagulase-negative staphylococci in milk from cows with clinical mastitis in Tunisia. Int. J. Antimicrob. Agents 52, 930–935. https://doi.org/10.1016/j.ijantimicag.2018.07.026

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Ding, L., 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25, 2283–2285. https://doi.org/10.1093/bioinformatics/btp373

Kocianova, S., Vuong, C., Yao, Y., Voyich, J.M., Fischer, E.R., DeLeo, F.R., Otto, M., 2005. Key role of poly-γ-dl-glutamic acid in immune evasion and virulence of Staphylococcus epidermidis. J Clin Invest 115, 688–694. https://doi.org/10.1172/JCI200523523

Köck, R., Schaumburg, F., Mellmann, A., Köksal, M., Jurke, A., Becker, K., Friedrich, A.W., 2013. Livestock-Associated Methicillin-Resistant Staphylococcus aureus (MRSA) as Causes of Human Infection and Colonization in Germany. PLOS ONE 8, e55040. https://doi.org/10.1371/journal.pone.0055040

Koreen, L., Ramaswamy, S.V., Graviss, E.A., Naidich, S., Musser, J.M., Kreiswirth, B.N., 2004. spa Typing Method for Discriminating among Staphylococcus aureus Isolates: Implications for Use of a Single Marker To Detect Genetic Micro- and Macrovariation. Journal of Clinical Microbiology 42, 792–799. https://doi.org/10.1128/JCM.42.2.792-799.2004

Kristian, S.A., Birkenstock, T.A., Sauder, U., Mack, D., Götz, F., Landmann, R., 2008. Biofilm formation induces C3a release and protects Staphylococcus epidermidis from IgG and complement deposition and from neutrophil-dependent killing. J. Infect. Dis. 197, 1028–1035. https://doi.org/10.1086/528992

Kuipers, A., Stapels, D.A.C., Weerwind, L.T., Ko, Y.-P., Ruyken, M., Lee, J.C., van Kessel, K.P.M., Rooijakkers, S.H.M., 2016. The Staphylococcus aureus polysaccharide capsule and Efb-dependent fibrinogen shield act in concert to protect against phagocytosis. Microbiology 162, 1185–1194. https://doi.org/10.1099/mic.0.000293

Kumar, R.M., Kaur, G., Kumar, N., Kumar, A., Singh, N.K., Bala, M., Kaur, N., Mayilraj, S., 2015. Taxonomic description and genome sequence of Salinicoccus sediminis spp.. nov., a halotolerant bacterium isolated from marine sediment. International Journal of Systematic and Evolutionary Microbiology 65, 3794–3799. https://doi.org/10.1099/ijsem.0.000495

Kwok, A.Y.C., 2003. Phylogenetic study of Staphylococcus and Macrococcus species based on partial hsp60 gene sequences. INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY 53, 87–92. https://doi.org/10.1099/ijs.0.02210-0

Kwong, J.C., McCallum, N., Sintchenko, V., Howden, B.P., 2015. Whole genome sequencing in clinical and public health microbiology. Pathology 47, 199–210. https://doi.org/10.1097/PAT.0000000000000235

LaBreck, P.T., Rice, G.K., Paskey, A.C., Elassal, E.M., Cer, R.Z., Law, N.N., Schlett, C.D., Bennett, J.W., Millar, E.V., Ellis, M.W., Hamilton, T., Bishop-Lilly, K.A., Merrell, D.S., 2018. Conjugative Transfer of a Novel Staphylococcal Plasmid

Encoding the Biocide Resistance Gene, qacA. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.02664

Lassen, S.B., Lomholt, H.B., Brüggemann, H., 2017. Complete Genome Sequence of a Staphylococcus epidermidis Strain with Exceptional Antimicrobial Activity. Genome Announc 5. https://doi.org/10.1128/genomeA.00004-17

Lavigne, J.-P., Espinal, P., Dunyach-Remy, C., Messad, N., Pantel, A., Sotto, A., 2013. Mass spectrometry: a revolution in clinical microbiology? Clin. Chem. Lab. Med. 51, 257–270. https://doi.org/10.1515/cclm-2012-0291

Lees, J.A., Harris, S.R., Tonkin-Hill, G., Gladstone, R.A., Lo, S.W., Weiser, J.N., Corander, J., Bentley, S.D., Croucher, N.J., 2019. Fast and flexible bacterial genomic epidemiology with PopPUNK. Genome Res. 29, 304–316. https://doi.org/10.1101/gr.241455.118

Letunic, I., Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44, W242-245. https://doi.org/10.1093/nar/gkw290

Li, D., Wu, C., Wang, Y., Fan, R., Schwarz, S., Zhang, S., 2015. Identification of Multiresistance Gene cfr in Methicillin-Resistant Staphylococcus aureus from Pigs: Plasmid Location and Integration into a Staphylococcal Cassette Chromosome mec Complex. Antimicrob Agents Chemother 59, 3641–3644. https://doi.org/10.1128/AAC.00500-15

Li, S., Skov, R.L., Han, X., Larsen, A.R., Larsen, J., Sørum, M., Wulf, M., Voss, A., Hiramatsu, K., Ito, T., 2011. Novel Types of Staphylococcal Cassette Chromosome mec Elements Identified in Clonal Complex 398 Methicillin-Resistant Staphylococcus aureus Strains. Antimicrobial Agents and Chemotherapy 55, 3046–3050. https://doi.org/10.1128/AAC.01475-10

Lilje, B., Rasmussen, R.V., Dahl, A., Stegger, M., Skov, R.L., Fowler, V.G., Ng, K.L., Kiil, K., Larsen, A.R., Petersen, A., Johansen, H.K., Schønheyder, H.C., Arpi, M., Rosenvinge, F.S., Korup, E., Høst, U., Hassager, C., Gill, S.U.A., Hansen, T.F., Johannesen, T.B., Smit, J., Søgaard, P., Skytt Andersen, P., Eske-Bruun, N., 2017. Whole-genome sequencing of bloodstream Staphylococcus aureus isolates does not distinguish bacteraemia from endocarditis. Microb Genom 3. https://doi.org/10.1099/mgen.0.000138

Lin, Y.-T., Tsai, J.-C., Chen, H.-J., Hung, W.-C., Hsueh, P.-R., Teng, L.-J., 2014. A Novel Staphylococcal Cassette Chromosomal Element, SCCfusC, Carrying fusC and speG in Fusidic Acid-Resistant Methicillin-Resistant Staphylococcus aureus. Antimicrob Agents Chemother 58, 1224–1227. https://doi.org/10.1128/AAC.01772-13

Lindsay, J.A., 2010. Genomic variation and evolution of Staphylococcus aureus. International Journal of Medical Microbiology 300, 98–103. https://doi.org/10.1016/j.ijmm.2009.08.013

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

Liu, G.Y., 2009. Molecular Pathogenesis of Staphylococcus aureus Infection. Pediatr Res 65, 71R-77R. https://doi.org/10.1203/PDR.0b013e31819dc44d

Ltd, B.P.G., 2016. Fosfomycin for UTIs. DTB 54, 114–117. https://doi.org/10.1136/dtb.2016.10.0429

Loncaric, I., Künzel, F., Klang, A., Wagner, R., Licka, T., Grunert, T., Feßler, A.T., Geier-Dömling, D., Rosengarten, R., Müller, E., Reissig, A., Spergser, J., Schwarz, S., Ehricht, R., Monecke, S., 2016. Carriage of meticillin-resistant

staphylococci between humans and animals on a small farm. Vet. Dermatol. 27, 191-e48. https://doi.org/10.1111/vde.12321

Lu, B., Leong, H.W., 2016. Computational methods for predicting genomic islands in microbial genomes. Computational and Structural Biotechnology Journal 14, 200–206. https://doi.org/10.1016/j.csbj.2016.05.001

Lutz, J.K., van Balen, J., Crawford, J.M., Wilkins, J.R., Lee, J., Nava-Hoet, R.C., Hoet, A.E., 2014. Methicillin-resistant Staphylococcus aureus in public transportation vehicles (buses): another piece to the epidemiologic puzzle. Am J Infect Control 42, 1285–1290. https://doi.org/10.1016/j.ajic.2014.08.016

Lymperopoulou, D.S., Coil, D.A., Schichnes, D., Lindow, S.E., Jospin, G., Eisen, J.A., Adams, R.I., 2017a. Draft genome sequences of eight bacteria isolated from the indoor environment: Staphylococcus capitis strain H36, S. capitis strain H65, S. cohnii strain H62, S. hominis strain H69, Microbacterium spp.. strain H83, Mycobacterium iranicum strain H39, Plantibacter spp.. strain H53, and Pseudomonas oryzihabitans strain H72. Stand Genomic Sci 12. https://doi.org/10.1186/s40793-017-0223-9

Lymperopoulou, D.S., Coil, D.A., Schichnes, D., Lindow, S.E., Jospin, G., Eisen, J.A., Adams, R.I., 2017b. Draft genome sequences of eight bacteria isolated from the indoor environment: Staphylococcus capitis strain H36, S. capitis strain H65, S. cohnii strain H62, S. hominis strain H69, Microbacterium spp.. strain H83, Mycobacterium iranicum strain H39, Plantibacter spp.. strain H53, and Pseudomonas oryzihabitans strain H72. Stand Genomic Sci 12. https://doi.org/10.1186/s40793-017-0223-9

MacLea, K.S., Trachtenberg, A.M., 2017. Complete Genome Sequence of Staphylococcus epidermidis ATCC 12228 Chromosome and Plasmids, Generated

by Long-Read Sequencing. Genome Announc 5. https://doi.org/10.1128/genomeA.00954-17

Madhusoodanan, J., Seo, K.S., Remortel, B., Park, J.Y., Hwang, S.Y., Fox, L.K., Park, Y.H., Deobald, C.F., Wang, D., Liu, S., Daugherty, S.C., Gill, A.L., Bohach, G.A., Gill, S.R., 2011. An Enterotoxin-Bearing Pathogenicity Island in Staphylococcus epidermidis ▿ . J Bacteriol 193, 1854–1862. https://doi.org/10.1128/JB.00162-10

Mafiz, A.I., Perera, L.N., He, Y., Zhang, W., Xiao, S., Hao, W., Sun, S., Zhou, K., Zhang, Y., 2018. Case study on the soil antibiotic resistome in an urban community garden. International Journal of Antimicrobial Agents 52, 241–250. https://doi.org/10.1016/j.ijantimicag.2018.05.016

Magaña-Lizárraga, J.A., Hernández-Peinado, J.V., Ahumada-Santos, Y.P., Parra-Unda, J.R., Uribe-Beltrán, M. de J., Gómez-Gil, B., Báez-Flores, M.E., 2017. Draft Genome Sequence of a Mexican Community-Associated Methicillin-Resistant Staphylococcus epidermidis Strain. Genome Announc 5. https://doi.org/10.1128/genomeA.01236-17

Malachowa, N., DeLeo, F.R., 2010. Mobile genetic elements of Staphylococcus aureus. Cell. Mol. Life Sci. 67, 3057–3071. https://doi.org/10.1007/s00018-010-0389-4

Maleki, A., Ghafourian, S., Taherikalani, M., Soroush, S., 2018. Alarming and threatening signals from health centers about multi drug resistance Staphylococcus haemolyticus. Infect Disord Drug Targets. https://doi.org/10.2174/1871526518666180911142806

Manara, S., Pasolli, E., Dolce, D., Ravenni, N., Campana, S., Armanini, F., Asnicar, F., Mengoni, A., Galli, L., Montagnani, C., Venturini, E., Rota-Stabelli, O., Grandi, G., Taccetti, G., Segata, N., 2018. Whole-genome epidemiology, characterisation,

and phylogenetic reconstruction of Staphylococcus aureus strains in a paediatric hospital. Genome Med 10, 82. https://doi.org/10.1186/s13073-018-0593-7

Maniv, I., Jiang, W., Bikard, D., Marraffini, L.A., 2016. Impact of Different Target Sequences on Type III CRISPR-Cas Immunity. Journal of bacteriology 198, 941–950. https://doi.org/10.1128/JB.00897-15

Maree, C.L., Daum, R.S., Boyle-Vavra, S., Matayoshi, K., Miller, L.G., 2007. Community-associated Methicillin-resistant Staphylococcus aureus Isolates and Healthcare-Associated Infections1. Emerg Infect Dis 13, 236–242. https://doi.org/10.3201/eid1302.060781

Martinez, J.L., Baquero, F., 2000. Mutation Frequencies and Antibiotic Resistance. Antimicrobial Agents and Chemotherapy 44, 1771–1777. https://doi.org/10.1128/AAC.44.7.1771-1777.2000

Mašlaňová, I., Wertheimer, Z., Sedláček, I., Švec, P., Indráková, A., Kovařovic, V., Schumann, P., Spröer, C., Králová, S., Šedo, O., Krištofová, L., Vrbovská, V., Füzik, T., Petráš, P., Zdráhal, Z., Ružičková, V., Doškař, J., Pantuček, R., 2018. Description and Comparative Genomics of Macrococcus caseolyticus subsp. hominis subsp. nov., Macrococcus goetzii spp.. nov., Macrococcus epidermidis spp.. nov., and Macrococcus bohemicus spp.. nov., Novel Macrococci From Human Clinical Material With Virulence Potential and Suspected Uptake of Foreign DNA by Natural Transformation. Front Microbiol 9. https://doi.org/10.3389/fmicb.2018.01178

Massimi, I., Park, E., Rice, K., Muller-Esterl, W., Sauder, D., McGavin, M.J., 2002. Identification of a novel maturation mechanism and restricted substrate specificity for the SspB cysteine protease of Staphylococcus aureus. J. Biol. Chem. 277, 41770–41777. https://doi.org/10.1074/jbc.M207162200

May, L., Klein, E.Y., Rothman, R.E., Laxminarayan, R., 2014. Trends in Antibiotic Resistance in Coagulase-Negative Staphylococci in the United States, 1999 to 2012. Antimicrob Agents Chemother 58, 1404–1409. https://doi.org/10.1128/AAC.01908-13

McNally, A., Oren, Y., Kelly, D., Pascoe, B., Dunn, S., Sreecharan, T., Vehkala, M., Välimäki, N., Prentice, M.B., Ashour, A., Avram, O., Pupko, T., Dobrindt, U., Literak, I., Guenther, S., Schaufler, K., Wieler, L.H., Zhiyong, Z., Sheppard, S.K., McInerney, J.O., Corander, J., 2016. Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. PLoS Genet 12. https://doi.org/10.1371/journal.pgen.1006280

Michael, C.A., Dominey-Howes, D., Labbate, M., 2014. The Antimicrobial Resistance Crisis: Causes, Consequences, and Management. Front Public Health 2. https://doi.org/10.3389/fpubh.2014.00145

Millan, A.S., 2018. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. Trends in Microbiology 26, 978–985. https://doi.org/10.1016/j.tim.2018.06.007

Miragaia, M., Carriço, J.A., Thomas, J.C., Couto, I., Enright, M.C., Lencastre, H. de, 2008. Comparison of Molecular Typing Methods for Characterization of Staphylococcus epidermidis: Proposal for Clone Definition. Journal of Clinical Microbiology 46, 118–129. https://doi.org/10.1128/JCM.01685-07

Miragaia, M., Lencastre, H. de, Perdreau-Remington, F., Chambers, H.F., Higashi, J., Sullam, P.M., Lin, J., Wong, K.I., King, K.A., Otto, M., Sensabaugh, G.F., Diep, B.A., 2009. Genetic Diversity of Arginine Catabolic Mobile Element in

Staphylococcus epidermidis. PLOS ONE 4, e7722. https://doi.org/10.1371/journal.pone.0007722

Miragaia, M., Thomas, J.C., Couto, I., Enright, M.C., de Lencastre, H., 2007. Inferring a population structure for Staphylococcus epidermidis from multilocus sequence typing data. J. Bacteriol. 189, 2540–2552. https://doi.org/10.1128/JB.01484-06

Mkrtchyan, H.V., Russell, C.A., Wang, N., Cutler, R.R., 2013. Could Public Restrooms Be an Environment for Bacterial Resistomes? PLOS ONE 8, e54223. https://doi.org/10.1371/journal.pone.0054223

Molloy, E.M., Casjens, S.R., Cox, C.L., Maxson, T., Ethridge, N.A., Margos, G., Fingerle, V., Mitchell, D.A., 2015. Identification of the minimal cytolytic unit for streptolysin S and an expansion of the toxin family. BMC Microbiology 15, 141. https://doi.org/10.1186/s12866-015-0464-y

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

Munita, J.M., Arias, C.A., 2016. Mechanisms of Antibiotic Resistance. Microbiol Spectr 4. https://doi.org/10.1128/microbiolspec.VMBF-0016-2015

Naimi, T.S., 2003. Comparison of Community- and Health Care–Associated Methicillin-Resistant Staphylococcus aureus Infection. JAMA 290, 2976. https://doi.org/10.1001/jama.290.22.2976

Nanoukon, C., Argemi, X., Sogbo, F., Orekan, J., Keller, D., Affolabi, D., Schramm, F., Riegel, P., Baba-Moussa, L., Prévost, G., 2017. Pathogenic features of clinically

significant coagulase-negative staphylococci in hospital and community infections in Benin. Int. J. Med. Microbiol. 307, 75–82. https://doi.org/10.1016/j.ijmm.2016.11.001

Naushad, S., Barkema, H.W., Luby, C., Condas, L.A., Nobrega, D.B., Carson, D.A., De, J.B., 2016. Comprehensive Phylogenetic Analysis of Bovine Non-aureus Staphylococci Species Based on Whole-Genome Sequencing. Front Microbiol 7, 1990–1990. https://doi.org/10.3389/fmicb.2016.01990

Nemoy, L.L., Kotetishvili, M., Tigno, J., Keefer-Norris, A., Harris, A.D., Perencevich, E.N., Johnson, J.A., Torpey, D., Sulakvelidze, A., Morris, J.G., Stine, O.C., 2005. Multilocus Sequence Typing versus Pulsed-Field Gel Electrophoresis for Characterization of Extended-Spectrum Beta-Lactamase-Producing Escherichia coli Isolates. J Clin Microbiol 43, 1776–1781. https://doi.org/10.1128/JCM.43.4.1776-1781.2005

Noto, M.J., Kreiswirth, B.N., Monk, A.B., Archer, G.L., 2008. Gene Acquisition at the Insertion Site for SCCmec, the Genomic Island Conferring Methicillin Resistance in Staphylococcus aureus. J Bacteriol 190, 1276–1283. https://doi.org/10.1128/JB.01128-07

O'Brien, F.G., Price, C., Grubb, W.B., Gustafson, J.E., 2002. Genetic characterization of the fusidic acid and cadmium resistance determinants of Staphylococcus aureus plasmid pUB101. J. Antimicrob. Chemother. 50, 313–321.

Oliveira, D.C., Tomasz, A., de Lencastre, H., 2002. Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant Staphylococcus aureus. Lancet Infect Dis 2, 180–189.

Oliveira, C.F. de, Cavanagh, J.P., Fredheim, E.G.A., Reiter, K.C., Rieger, A., Klingenberg, C., d'Azevedo, P.A., Sollid, J.E., Oliveira, C.F. de, Cavanagh, J.P.,

Fredheim, E.G.A., Reiter, K.C., Rieger, A., Klingenberg, C., d'Azevedo, P.A., Sollid, J.E., 2016. Coagulase-negative staphylococci in Southern Brazil: looking toward its high diversity. Revista da Sociedade Brasileira de Medicina Tropical 49, 292–299. https://doi.org/10.1590/0037-8682-0015-2016

O'neill, J.f, 2016 Tackling Drug-Resistant Infections Globally: final report and recommendations.

O'Riordan, K., Lee, J.C., 2004. Staphylococcus aureus Capsular Polysaccharides. Clin Microbiol Rev 17, 218–234. https://doi.org/10.1128/CMR.17.1.218-234.2004

Otto, M., 2012. Molecular basis of Staphylococcus epidermidis infections. Semin Immunopathol 34, 201–214. https://doi.org/10.1007/s00281-011-0296-2

Otto, M., 2009. Staphylococcus epidermidis – the "accidental" pathogen. Nat Rev Microbiol 7, 555–567. https://doi.org/10.1038/nrmicro2182

Otto, M., 2008. Staphylococcal Biofilms. Curr Top Microbiol Immunol 322, 207–228.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421

Panda, S., Singh, D.V., 2016. Whole-Genome Sequences of Staphylococcus haemolyticus Isolated from Infected Eyes and Healthy Conjunctiva in Bhubaneswar, India. Genome Announc 4. https://doi.org/10.1128/genomeA.00099-16

Park, N., 2017. Lower layer Super Output Area population estimates (supporting information) - Office for National Statistics [WWW Document]. Office for National Statistics. URL https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigratio

n/populationestimates/datasets/lowersuperoutputareamidyearpopulationestimates (accessed 12.4.18).

Park, S., Kelley, K.A., Vinogradov, E., Solinga, R., Weidenmaier, C., Misawa, Y., Lee, J.C., 2010. Characterization of the Structure and Biological Functions of a Capsular Polysaccharide Produced by Staphylococcus saprophyticus. Journal of Bacteriology 192, 4618–4626. https://doi.org/10.1128/JB.00104-10

Paterson, G.K., Harrison, E.M., Holmes, M.A., 2014. The emergence of mecC methicillin-resistant Staphylococcus aureus. Trends Microbiol 22, 42–47. https://doi.org/10.1016/j.tim.2013.11.003

Pedroso, S.H.S.P., Sandes, S.H.C., Filho, R.A.T., Nunes, A.C., Serufo, J.C., Farias, L.M., Carvalho, M.A.R., Bomfim, M.R.Q., Santos, S.G., 2018. Coagulase-Negative Staphylococci Isolated from Human Bloodstream Infections Showed Multidrug Resistance Profile. Microb. Drug Resist. 24, 635–647. https://doi.org/10.1089/mdr.2017.0309

Pereira, P.M.A., Binatti, V.B., Sued, B.P.R., Ramos, J.N., Peixoto, R.S., Simões, C., de Castro, E.A., Duarte, J.L.M.B., Vieira, V.V., Hirata, R., Santos, K.R.N., Mattos-Guaraldi, A.L., Pereira, J.A.A., 2014. Staphylococcus haemolyticus disseminated among neonates with bacteremia in a neonatal intensive care unit in Rio de Janeiro, Brazil. Diagn. Microbiol. Infect. Dis. 78, 85–92. https://doi.org/10.1016/j.diagmicrobio.2013.06.026

Peters, G., Locci, R., Pulverer, G., 1982. Adherence and growth of coagulase-negative staphylococci on surfaces of intravenous catheters. J. Infect. Dis. 146, 479–482.

Petersson, A.C., Olsson-Liljequist, B., Miörner, H., Hæggman, S., 2010. Evaluating the usefulness of spa typing, in comparison with pulsed-field gel electrophoresis, for epidemiological typing of methicillin-resistant Staphylococcus aureus in a low-

prevalence region in Sweden 2000–2004. Clinical Microbiology and Infection 16, 456–462. https://doi.org/10.1111/j.1469-0691.2009.02881.x

Petinaki, E., Arvaniti, A., Dimitracopoulos, G., Spiliopoulou, I., 2001. Detection of mecA, mecR1 and mecI genes among clinical isolates of methicillin-resistant staphylococci by combined polymerase chain reactions. J. Antimicrob. Chemother. 47, 297–304.

Peton, V., Le Loir, Y., 2014. Staphylococcus aureus in veterinary medicine. Infection, Genetics and Evolution 21, 602–615. https://doi.org/10.1016/j.meegid.2013.08.011

Planet, P.J., LaRussa, S.J., Dana, A., Smith, H., Xu, A., Ryan, C., Uhlemann, A.-C., Boundy, S., Goldberg, J., Narechania, A., Kulkarni, R., Ratner, A.J., Geoghegan, J.A., Kolokotronis, S.-O., Prince, A., 2013. Emergence of the Epidemic Methicillin-Resistant Staphylococcus aureus Strain USA300 Coincides with Horizontal Transfer of the Arginine Catabolic Mobile Element and speG-mediated Adaptations for Survival on Skin. mBio 4, e00889-13. https://doi.org/10.1128/mBio.00889-13

Polz, M.F., Alm, E.J., Hanage, W.P., 2013. Horizontal Gene Transfer and the Evolution of Bacterial and Archaeal Population Structure. Trends Genet 29, 170–175. https://doi.org/10.1016/j.tig.2012.12.006

Prado, M.R., Blandón, L.M., Vandenberghe, L.P.S., Rodrigues, C., Castro, G.R., Thomaz-Soccol, V., Soccol, C.R., 2015. Milk kefir: composition, microbial cultures, biological activities, and related products. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.01177

Premkrishnan, B.N.V., Junqueira, A.C.M., Uchida, A., Purbojati, R.W., Houghton, J.N.I., Chénard, C., Wong, A., Kolundžija, S., Clare, M.E., Kushwaha, K.K., Panicker,

D., Putra, A., Gaultier, N.E., Heinle, C.E., Vettath, V.K., Drautz-Moses, D.I., Schuster, S.C., 2018. Complete Genome Sequence of Staphylococcus haemolyticus Type Strain SGAir0252. Genome Announc 6. https://doi.org/10.1128/genomeA.00229-18

Pruitt, K.D., Tatusova, T., Maglott, D.R., 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35, D61–D65. https://doi.org/10.1093/nar/gkl842

Purves, J., Thomas, J., Riboldi, G.P., Zapotoczna, M., Tarrant, E., Andrew, P.W., Londoño, A., Planet, P.J., Geoghegan, J.A., Waldron, K.J., Morrissey, J.A., 2018. A horizontally gene transferred copper resistance locus confers hyper-resistance to antibacterial copper toxicity and enables survival of community acquired methicillin resistant Staphylococcus aureus USA300 in macrophages. Environ Microbiol 20, 1576–1589. https://doi.org/10.1111/1462-2920.14088

Queck, S.Y., Khan, B.A., Wang, R., Bach, T.-H.L., Kretschmer, D., Chen, L., Kreiswirth, B.N., Peschel, A., DeLeo, F.R., Otto, M., 2009. Mobile Genetic Element-Encoded Cytolysin Connects Virulence to Methicillin Resistance in MRSA. PLOS Pathogens 5, e1000533. https://doi.org/10.1371/journal.ppat.1000533

Rahman, B., Ownagh, A., Mardani, K., Farrokhi Ardebili, F., 2016. Prevalence and molecular characterization of staphylococci isolated from sheep with subclinical mastitis in West-Azerbaijan province, Iran. Vet Res Forum 7, 155–162.

Ramisetty, B.C.M., Sudhakari, P.A., 2019. Bacterial 'Grounded' Prophages: Hotspots for Genetic Renovation and Innovation. Front Genet 10. https://doi.org/10.3389/fgene.2019.00065

Ravenhall, M., Škunca, N., Lassalle, F., Dessimoz, C., 2015. Inferring Horizontal Gene Transfer. PLOS Computational Biology 11, e1004095. https://doi.org/10.1371/journal.pcbi.1004095Ra

Regué, M., Hita, B., Piqué, N., Izquierdo, L., Merino, S., Fresno, S., Benedí, V.J., Tomás, J.M., 2004. A Gene, uge, Is Essential for Klebsiella pneumoniae Virulence. Infect Immun 72, 54–61. https://doi.org/10.1128/IAI.72.1.54-61.2004

Rivera-Perez, J.I., Santiago-Rodriguez, T.M., Toranzos, G.A., 2016. Paleomicrobiology: a Snapshot of Ancient Microbes and Approaches to Forensic Microbiology. Microbiol Spectr 4. https://doi.org/10.1128/microbiolspec.EMF-0006-2015

Roberts, M.C., Soge, O.O., No, D., 2013. Comparison of Multi-Drug Resistant Environmental Methicillin-Resistant Staphylococcus aureus Isolated from Recreational Beaches and High Touch Surfaces in Built Environments. Front. Microbiol. 4. https://doi.org/10.3389/fmicb.2013.00074

Ross, A., Ward, S., Hyman, P., 2016. More Is Better: Selecting for Broad Host Range Bacteriophages. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.01352

Rumpel, S., Razeto, A., Pillar, C.M., Vijayan, V., Taylor, A., Giller, K., Gilmore, M.S., Becker, S., Zweckstetter, M., 2004. Structure and DNA-binding properties of the cytolysin regulator CylR2 from Enterococcus faecalis. EMBO J 23, 3632–3642. https://doi.org/10.1038/sj.emboj.7600367

Ruzauskas, M., Siugzdiniene, R., Klimiene, I., Virgailis, M., Mockeliunas, R., Vaskeviciute, L., Zienius, D., 2014. Prevalence of methicillin-resistant Staphylococcus haemolyticus in companion animals: a cross-sectional study. Ann. Clin. Microbiol. Antimicrob. 13, 56. https://doi.org/10.1186/s12941-014-0056-y

Sabat, A.J., Budimir, A., Nashev, D., Sá-Leão, R., Dijl, J.M. van, Laurent, F., Grundmann, H., Friedrich, A.W., Markers (ESGEM), on behalf of the E.S.G. of E., 2013.

Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Eurosurveillance 18, 20380. https://doi.org/10.2807/ese.18.04.20380-en

Saeb, A.T.M., 2018. Current Bioinformatics resources in combating infectious diseases. Bioinformation 14, 31–35. https://doi.org/10.6026/97320630014031

Sasaki, T., Tsubakishita, S., Tanaka, Y., Sakusabe, A., Ohtsuka, M., Hirotaki, S., Kawakami, T., Fukata, T., Hiramatsu, K., 2010. Multiplex-PCR Method for Species Identification of Coagulase-Positive Staphylococci. J Clin Microbiol 48, 765–769. https://doi.org/10.1128/JCM.01232-09

Savijoki, K., Iivanainen, A., Siljamäki, P., Laine, P.K., Paulin, L., Karonen, T., Pyörälä, S., Kankainen, M., Nyman, T.A., Salomäki, T., Koskinen, P., Holm, L., Simojoki, H., Taponen, S., Sukura, A., Kalkkinen, N., Auvinen, P., Varmanen, P., 2014. Genomics and Proteomics Provide New Insight into the Commensal and Pathogenic Lifestyles of Bovine- and Human-Associated Staphylococcus epidermidis Strains. J. Proteome Res. 13, 3748–3762. https://doi.org/10.1021/pr500322d

Scarff, J.M., Seldina, Y.I., Vergis, J.M., Ventura, C.L., O'Brien, A.D., 2018. Expression and contribution to virulence of each polysaccharide capsule of Bacillus cereus strain G9241. PLOS ONE 13, e0202701. https://doi.org/10.1371/journal.pone.0202701

Schoenfelder, S.M.K., Dong, Y., Feßler, A.T., Schwarz, S., Schoen, C., Köck, R., Ziebuhr, W., 2017. Antibiotic resistance profiles of coagulase-negative staphylococci in livestock environments. Vet. Microbiol. 200, 79–87. https://doi.org/10.1016/j.vetmic.2016.04.019

Schubert, S., Kostrzewa, M., 2017. MALDI-TOF MS in the Microbiology Laboratory: Current Trends. Curr Issues Mol Biol 23, 17–20. https://doi.org/10.21775/cimb.023.017

Schwarz, S., Kehrenberg, C., Ojo, K.K., 2002. Staphylococcus sciuri Gene erm(33), Encoding Inducible Resistance to Macrolides, Lincosamides, and Streptogramin B Antibiotics, Is a Product of Recombination between erm(C) and erm(A). Antimicrob Agents Chemother 46, 3621–3623. https://doi.org/10.1128/AAC.46.11.3621-3623.2002

Scybert, S., Pechous, R., Sitthisak, S., Nadakavukaren, M.J., Wilkinson, B.J., Jayaswal, R.K., 2003. NaCl-sensitive mutant of *Staphylococcus aureus* has a Tn917- *lacZ* insertion in its *ars* operon. FEMS Microbiology Letters 222, 171–176. https://doi.org/10.1016/S0378-1097(03)00312-4

Segerman, B., 2012. The genetic integrity of bacterial species: the core genome and the accessory genome, two different stories. Front Cell Infect Microbiol 2. https://doi.org/10.3389/fcimb.2012.00116

Seng, R., Kitti, T., Thummeepak, R., Kongthai, P., Leungtongkam, U., Wannalerdsakun, S., Sitthisak, S., 2017a. Biofilm formation of methicillin-resistant coagulase negative staphylococci (MR-CoNS) isolated from community and hospital environments. PLOS ONE 12, e0184172. https://doi.org/10.1371/journal.pone.0184172

Seng, R., Leungtongkam, U., Thummeepak, R., Chatdumrong, W., Sitthisak, S., 2017b. High prevalence of methicillin-resistant coagulase-negative staphylococci isolated from a university environment in Thailand. Int. Microbiol. 20, 65–73. https://doi.org/10.2436/20.1501.01.286

Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W., Vaishampayan, P., 2018. Development of a Custom MALDI-TOF MS Database for Species-Level Identification of Bacterial Isolates Collected From Spacecraft and Associated Surfaces. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.00780

Shahbazian, J.H., Hahn, P.D., Ludwig, S., Ferguson, J., Baron, P., Christ, A., Spicer, K., Tolomeo, P., Torrie, A.M., Bilker, W.B., Cluzet, V.C., Hu, B., Julian, K., Nachamkin, I., Rankin, S.C., Morris, D.O., Lautenbach, E., Davis, M.F., 2017. Multidrug and mupirocin resistance in environmental methicillin-resistant Staphylococcus aureus (MRSA) collected from the homes of people diagnosed with a community-onset (CO-) MRSA infection. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.01369-17

Shallcross, L.J., Davies, D.S.C., 2014. Antibiotic overuse: a key driver of antimicrobial resistance. Br J Gen Pract 64, 604–605. https://doi.org/10.3399/bjgp14X682561

Shankar, N., Coburn, P., Pillar, C., Haas, W., Gilmore, M., 2004. Enterococcal cytolysin: activities and association with other virulence traits in a pathogenicity island. Int. J. Med. Microbiol. 293, 609–618. https://doi.org/10.1078/1438-4221-00301

Sheng, H., Nakamura, K., Kanno, T., Sasaki, K., Niwano, Y., 2015. Microbicidal Activity of Artificially Generated Hydroxyl Radicals, in: Sasaki, K., Suzuki, O., Takahashi, N. (Eds.), Interface Oral Health Science 2014. Springer Japan, pp. 203–215.

Shittu, A., Lin, J., Morrison, D., Kolawole, D., 2006. Identification and molecular characterization of mannitol salt positive, coagulase-negative staphylococci from nasal samples of medical personnel and students. J. Med. Microbiol. 55, 317–324. https://doi.org/10.1099/jmm.0.46072-0

Shore, A., Rossney, A.S., Keane, C.T., Enright, M.C., Coleman, D.C., 2005. Seven novel variants of the staphylococcal chromosomal cassette mec in methicillin-resistant

Staphylococcus aureus isolates from Ireland. Antimicrob. Agents Chemother. 49, 2070–2083. https://doi.org/10.1128/AAC.49.5.2070-2083.2005

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

Silva, Patricia Vollu, Cruz, R.S., Keim, L.S., Paula, G.R. de, Carvalho, B.T.F., Coelho, L.R., Carvalho, M.C. da S., Rosa, J.M.C. da, Figueiredo, A.M.S., Teixeira, L.A., 2013. The antimicrobial susceptibility, biofilm formation and genotypic profiles of Staphylococcus haemolyticus from bloodstream infections. Memórias do Instituto Oswaldo Cruz 108, 812–813. https://doi.org/10.1590/0074-02761080620130022

Simeone, R., Bottai, D., Brosch, R., 2009. ESX/type VII secretion systems and their role in host–pathogen interaction. Current Opinion in Microbiology, Host–Microbe Interactions: Bacteria 12, 4–10. https://doi.org/10.1016/j.mib.2008.11.003

Singh, J.K., Adams, F.G., Brown, M.H., 2019. Diversity and Function of Capsular Polysaccharide in Acinetobacter baumannii. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.03301

Singhal, N., Kumar, M., Kanaujia, P.K., Virdi, J.S., 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol 6, 791. https://doi.org/10.3389/fmicb.2015.00791

Soge, O.O., Meschke, J.S., No, D.B., Roberts, M.C., 2009. Characterization of methicillin-resistant Staphylococcus aureus and methicillin-resistant coagulase-

negative Staphylococcus spp.. isolated from US West Coast public marine beaches. J. Antimicrob. Chemother. 64, 1148–1155. https://doi.org/10.1093/jac/dkp368

Sollid, J.U.E., Furberg, A.S., Hanssen, A.M., Johannessen, M., 2014. Staphylococcus aureus: determinants of human carriage. Infect. Genet. Evol. 21, 531–541. https://doi.org/10.1016/j.meegid.2013.03.02

Soroush, S., Jabalameli, F., Taherikalani, M., Eslampour, M.A., Beigverdi, R., Emaneini, M., 2017. Characterization of biofilm formation, antimicrobial resistance, and staphylococcal cassette chromosome mec analysis of methicillin resistant Staphylococcus hominis from blood cultures of children. Rev. Soc. Bras. Med. Trop. 50, 329–333. https://doi.org/10.1590/0037-8682-0384-2016

Spanu, T., Sanguinetti, M., Ciccaglione, D., D'Inzeo, T., Romano, L., Leone, F., Fadda, G., 2003. Use of the VITEK 2 System for Rapid Identification of Clinical Isolates of Staphylococci from Bloodstream Infections. J Clin Microbiol 41, 4259–4263. https://doi.org/10.1128/JCM.41.9.4259-4263.2003

Stähler, F.N., Odenbreit, S., Haas, R., Wilrich, J., Vliet, A.H.M.V., Kusters, J.G., Kist, M., Bereswill, S., 2006. The Novel Helicobacter pylori CznABC Metal Efflux Pump Is Required for Cadmium, Zinc, and Nickel Resistance, Urease Modulation, and Gastric Colonization. Infect Immun 74, 3845–3852. https://doi.org/10.1128/IAI.02025-05

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Staphylococcus cohnii (ID 24233) - Genome - NCBI [WWW Document], n.d. URL https://www.ncbi.nlm.nih.gov/genome/24233?genome_assembly_id=366618 (accessed 12.3.18)stam

Staphylococcus hominis subsp. hominis (ID 2014) - Genome - NCBI [WWW Document], n.d. URL https://www.ncbi.nlm.nih.gov/genome/2014?genome_assembly_id=358743 (accessed 12.3.18).

Staphylococcus sciuri (ID 10782) - Genome - NCBI [WWW Document], n.d. URL https://www.ncbi.nlm.nih.gov/genome/10782?genome_assembly_id=366045 (accessed 12.3.18).

Stapleton, P.D., Taylor, P.W., 2002. Methicillin resistance in Staphylococcus aureus. Sci Prog 85, 57–72.

Stefani, S., Varaldo, P.E., 2003. Epidemiology of methicillin-resistant staphylococci in Europe. Clinical Microbiology and Infection 9, 1179–1186. https://doi.org/10.1111/j.1469-0691.2003.00698.x

Sung, J.M.-L., Lloyd, D.H., Lindsay, J.A., 2008. Staphylococcus aureus host specificity: comparative genomics of human versus animal isolates by multi-strain microarray. Microbiology 154, 1949–1959. https://doi.org/10.1099/mic.0.2007/015289-0

Suzuki, H., Lefébure, T., Bitar, P.P., Stanhope, M.J., 2012. Comparative genomic analysis of the genus Staphylococcus including Staphylococcus aureus and its newly described sister species Staphylococcus simiae. BMC Genomics 13, 38. https://doi.org/10.1186/1471-2164-13-38

Sweeney, M.T., Lubbers, B.V., Schwarz, S., Watts, J.L., 2018. Applying definitions for multidrug resistance, extensive drug resistance and pandrug resistance to

clinically significant livestock and companion animal bacterial pathogens. J. Antimicrob. Chemother. 73, 1460–1463. https://doi.org/10.1093/jac/dky043

Szczuka, E., Krajewska, M., Lijewska, D., Bosacka, K., Kaznowski, A., 2016. Diversity of staphylococcal cassette chromosome mec elements in nosocomial multiresistant Staphylococcus haemolyticus isolates. J Appl Genet 57, 543–547. https://doi.org/10.1007/s13353-016-0346-5

Szczuka, E., Krzymińska, S., Bogucka, N., Kaznowski, A., 2018. Multifactorial mechanisms of the pathogenesis of methicillin-resistant Staphylococcus hominis isolated from bloodstream infections. Antonie Van Leeuwenhoek 111, 1259–1265. https://doi.org/10.1007/s10482-017-1007-3

Szczuka, E., Makowska, N., Bosacka, K., Słotwińska, A., Kaznowski, A., 2016. Molecular basis of resistance to macrolides, lincosamides and streptogramins in Staphylococcus hominis strains isolated from clinical specimens. Folia Microbiol (Praha) 61, 143–147. https://doi.org/10.1007/s12223-015-0419-6

Szczuka, E., Telega, K., Kaznowski, A., 2015. Biofilm formation by Staphylococcus hominis strains isolated from human clinical specimens. Folia Microbiol. (Praha) 60, 1–5. https://doi.org/10.1007/s12223-014-0332-4

Takeuchi, F., Watanabe, S., Baba, T., Yuzawa, H., Ito, T., Morimoto, Y., Kuroda, M., Cui, L., Takahashi, M., Ankai, A., Baba, S., Fukui, S., Lee, J.C., Hiramatsu, K., 2005. Whole-Genome Sequencing of Staphylococcus haemolyticus Uncovers the Extreme Plasticity of Its Genome and the Evolution of Human-Colonizing Staphylococcal Species. J Bacteriol 187, 7292–7308. https://doi.org/10.1128/JB.187.21.7292-7308.2005

Tammelin, A., Klötz, F., Hambraeus, A., Ståhle, E., Ransjö, U., 2003. Nasal and hand carriage of Staphylococcus aureus in staff at a Department for Thoracic and

Cardiovascular Surgery: endogenous or exogenous source? Infect Control Hosp Epidemiol 24, 686–689. https://doi.org/10.1086/502277

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4, 41. https://doi.org/10.1186/1471-2105-4-41

Tchoupa, A.K., Jones, R.A., Kuroki, A., Alam, M.T., Perrier, S., Chen, Y., Unnikrishnan, M., 2019. The type VII secretion system protects Staphylococcus aureus against antimicrobial host fatty acids. bioRxiv 572172. https://doi.org/10.1101/572172

Ternes, Y.M., Lamaro-Cardoso, J., André, M.C.P., Pessoa, V.P., Vieira, M.A. da S., Minamisava, R., Andrade, A.L., Kipnis, A., 2013. Molecular epidemiology of coagulase-negative Staphylococcus carriage in neonates admitted to an intensive care unit in Brazil. BMC Infect Dis 13, 572. https://doi.org/10.1186/1471-2334-13-572.

Thomas, J.C., Robinson, D.A., 2014. Multilocus sequence typing of Staphylococcus epidermidis. Methods Mol. Biol. 1106, 61–69. https://doi.org/10.1007/978-1-62703-736-5_5

Tong, S.Y.C., Davis, J.S., Eichenberger, E., Holland, T.L., Fowler, V.G., 2015. Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. Clin Microbiol Rev 28, 603–661. https://doi.org/10.1128/CMR.00134-14

Totake, K., Kobayashi, N., Odaka, T., 1998. Trimethoprim resistance and susceptibility genes in Staphylococcus epidermidis. Microbiol. Immunol. 42, 497–501. https://doi.org/10.1111/j.1348-0421.1998.tb02315.x

Trampuz, A., Zimmerli, W., 2005. Prosthetic joint infections: update in diagnosis and treatment. Swiss Med. Wkly. 135, 243–251.

Tsoi, H.-W., Tse, H., 2011. Staphylococcus lugdunensis Is the Likely Origin of the Ornithine Decarboxylase Operon in Staphylococcus epidermidis 2015. Appl Environ Microbiol 77, 392–393. https://doi.org/10.1128/AEM.02213-10

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., Hiramatsu, K., 2010. Origin and Molecular Evolution of the Determinant of Methicillin Resistance in Staphylococci. Antimicrob Agents Chemother 54, 4352–4359. https://doi.org/10.1128/AAC.00356-10

Tulinski, P., Fluit, A.C., Wagenaar, J.A., Mevius, D., Vijver, L. van de, Duim, B., 2012. Methicillin-Resistant Coagulase-Negative Staphylococci on Pig Farms as a Reservoir of Heterogeneous Staphylococcal Cassette Chromosome mec Elements. Appl. Environ. Microbiol. 78, 299–304. https://doi.org/10.1128/AEM.05594-11

Twine, S.M., Vinogradov, E., Lindgren, H., Sjostedt, A., Conlan, J.W., 2012. Roles for wbtC, wbtI, and kdtA Genes in Lipopolysaccharide Biosynthesis, Protein Glycosylation, Virulence, and Immunogenicity in Francisella tularensis2 Strain SCHU S4. Pathogens 1, 12–29. https://doi.org/10.3390/pathogens1010012

van Schaik, W., 2015. The human gut resistome. Philos Trans R Soc Lond B Biol Sci 370. https://doi.org/10.1098/rstb.2014.0087

Van Tyne, D., Martin, M.J., Gilmore, M.S., 2013. Structure, Function, and Biology of the Enterococcus faecalis Cytolysin. Toxins 5, 895–911. https://doi.org/10.3390/toxins5050895

Vandecandelaere, I., Van, F.N., Deforce, D., Nelis, H.J., Coenye, T., 2014. Draft Genome Sequence of Methicillin-Resistant Staphylococcus epidermidis Strain ET-024,

Isolated from an Endotracheal Tube Biofilm of a Mechanically Ventilated Patient. Genome Announc 2. https://doi.org/10.1128/genomeA.00527-14

Vanderhaeghen, W., Hermans, K., Haesebrouck, F., Butaye, P., 2010. Methicillin-resistant Staphylococcus aureus (MRSA) in food production animals. Epidemiology and Infection 138, 606. https://doi.org/10.1017/S0950268809991567

Vázquez-Boland, J.A., Kuhn, M., Berche, P., Chakraborty, T., Domínguez-Bernal, G., Goebel, W., González-Zorn, B., Wehland, J., Kreft, J., 2001. Listeria Pathogenesis and Molecular Virulence Determinants. Clin Microbiol Rev 14, 584–640. https://doi.org/10.1128/CMR.14.3.584-640.2001

Velasco, V., Buyukcangaz, E., Sherwood, J.S., Stepan, R.M., Koslofsky, R.J., Logue, C.M., 2015. Characterization of Staphylococcus aureus from Humans and a Comparison with İsolates of Animal Origin, in North Dakota, United States. PLoS One 10. https://doi.org/10.1371/journal.pone.0140497

Veloo, A.C.M., Jean-Pierre, H., Justesen, U.S., Morris, T., Urban, E., Wybo, I., Shah, H.N., Friedrich, A.W., ENRIA workgroup, Morris, T., Shah, H.N., Jean-Pierre, H., Justesen, U.S., Nagy, E., Urban, E., Kostrzewa, M., Veloo, A., Friedrich, A.W., 2017. A multi-center ring trial for the identification of anaerobic bacteria using MALDI-TOF MS. Anaerobe 48, 94–97. https://doi.org/10.1016/j.anaerobe.2017.07.004

Ventola, C.L., 2015. The Antibiotic Resistance Crisis. P T 40, 277–283.

Vernozy-Rozand, C., Mazuy, C., Meugnier, H., Bes, M., Lasne, Y., Fiedler, F., Etienne, J., Freney, J., 2000. Staphylococcus fleurettii spp.. nov., isolated from goat's milk cheeses. Int. J. Syst. Evol. Microbiol. 50 Pt 4, 1521–1527. https://doi.org/10.1099/00207713-50-4-1521

Voineagu, L., Braga, V., Botnarciuc, M., Barbu, A., Tataru, M., 2012. Emergence of Staphylococcus hominis Strains in General Infections. ARS Medica Tomitana 18, 80–82. https://doi.org/10.2478/v10307-012-0016-8

von Wintersdorff, C.J.H., Penders, J., van Niekerk, J.M., Mills, N.D., Majumder, S., van Alphen, L.B., Savelkoul, P.H.M., Wolffs, P.F.G., 2016. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.00173

Waack, S., Keller, O., Asper, R., Brodag, T., Damm, C., Fricke, W.F., Surovcik, K., Meinicke, P., Merkl, R., 2006. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics 7, 142. https://doi.org/10.1186/1471-2105-7-142

Walsh, T.R., 2018. A one-health approach to antimicrobial resistance. Nat Microbiol 3, 854–855. https://doi.org/10.1038/s41564-018-0208-5

Wang, C., Cai, P., Zhan, Q., Mi, Z., Huang, Z., Chen, G., 2008. Distribution of antiseptic-resistance genes qacA/B in clinical isolates of meticillin-resistant Staphylococcus aureus in China. J. Hosp. Infect. 69, 393–394. https://doi.org/10.1016/j.jhin.2008.05.009

Wang, J., Kuenzel, S., Baines, J.F., 2014. Draft Genome Sequences of 11 Staphylococcus epidermidis Strains Isolated from Wild Mouse Species. Genome Announc 2. https://doi.org/10.1128/genomeA.01148-13

Wang, W., Zheng, S.-S., Sharshov, K., Sun, H., Wu, X.-Q., Yang, F., Wang, X.-L., Li, L.-X., 2017. Draft Genome Sequence of Staphylococcus hominis BHG17 Isolated from Wild Bar-Headed Goose (Anser indicus) Feces. Genome Announc 5. https://doi.org/10.1128/genomeA.01552-16

Wang, X., Liu, Q., Zhang, H., Li, X., Huang, W., Fu, Q., Li, M., 2018. Molecular Characteristics of Community-Associated Staphylococcus aureus Isolates From Pediatric Patients With Bloodstream Infections Between 2012 and 2017 in Shanghai, China. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.01211

Waters, A.E., Contente-Cuomo, T., Buchhagen, J., Liu, C.M., Watson, L., Pearce, K., Foster, J.T., Bowers, J., Driebe, E.M., Engelthaler, D.M., Keim, P.S., Price, L.B., 2011. Multidrug-Resistant Staphylococcus aureus in US Meat and Poultry. Clin Infect Dis 52, 1227–1230. https://doi.org/10.1093/cid/cir181

Webber, M.A., Piddock, L.J.V., 2003. The importance of efflux pumps in bacterial antibiotic resistance. J Antimicrob Chemother 51, 9–11. https://doi.org/10.1093/jac/dkg050

Weese, J.S., 2010. Methicillin-resistant Staphylococcus aureus in animals. ILAR J 51, 233–244. https://doi.org/10.1093/ilar.51.3.233

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173, 697–703.

Widerström, M., McCullough, C.A., Coombs, G.W., Monsen, T., Christiansen, K.J., 2012. A Multidrug-Resistant Staphylococcus epidermidis Clone (ST2) Is an Ongoing Cause of Hospital-Acquired Infection in a Western Australian Hospital. Journal of Clinical Microbiology 50, 2147–2151. https://doi.org/10.1128/JCM.06456-11

Wojtyczka, R.D., Orlewska, K., Kępa, M., Idzik, D., Dziedzic, A., Mularz, T., Krawczyk, M., Miklasińska, M., Wąsik, T.J., 2014. Biofilm Formation and Antimicrobial Susceptibility of Staphylococcus epidermidis Strains from a Hospital Environment. Int J Environ Res Public Health 11, 4619–4633. https://doi.org/10.3390/ijerph110504619

Wos-Oxley, M.L., Plumeier, I., von Eiff, C., Taudien, S., Platzer, M., Vilchez-Vargas, R., Becker, K., Pieper, D.H., 2010. A poke into the diversity and associations within human anterior nare microbial communities. ISME J 4, 839–851. https://doi.org/10.1038/ismej.2010.15

Wright, G.D., 2007. The antibiotic resistome: the nexus of chemical and genetic diversity. Nature Reviews Microbiology 5, 175–186. https://doi.org/10.1038/nrmicro1614

Wu, S., Zhu, Z., Fu, L., Niu, B., Li, W., 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics 12, 444. https://doi.org/10.1186/1471-2164-12-444

Xu, Z., Misra, R., Jamrozy, D., Paterson, G.K., Cutler, R.R., Holmes, M.A., Gharbia, S., Mkrtchyan, H.V., 2018a. Whole Genome Sequence and Comparative Genomics Analysis of Multi-drug Resistant Environmental Staphylococcus epidermidis ST59. G3: Genes, Genomes, Genetics 8, 2225–2230. https://doi.org/10.1534/g3.118.200314

Xu, Z., Mkrtchyan, H.V., Cutler, R.R., 2015. Antibiotic resistance and mecA characterization of coagulase-negative staphylococci isolated from three hotels in London, UK. Front Microbiol 6. https://doi.org/10.3389/fmicb.2015.00947

Xu, Z., Shah, H.N., Misra, R., Chen, J., Zhang, W., Liu, Y., Cutler, R.R., Mkrtchyan, H.V., 2018b. The prevalence, antibiotic resistance and mecA characterization of coagulase negative staphylococci recovered from non-healthcare settings in London, UK. Antimicrobial Resistance & Infection Control 7, 73. https://doi.org/10.1186/s13756-018-0367-4

Xu, Z., Zhang, H., Sun, X., Liu, Y., Yan, W., Xun, W., Shen, Q., Zhang, R., 2019. Bacillus velezensis Wall Teichoic Acids Are Required for Biofilm Formation and

Root Colonization. Appl. Environ. Microbiol. 85, e02116-18. https://doi.org/10.1128/AEM.02116-18

Yamaguchi, A., Shiina, Y., Fujihira, E., Sawai, T., Noguchi, N., Sasatsu, M., 1995. The tetracycline efflux protein encoded by the tet(K) gene from Staphylococcus aureus is a metal-tetracycline/H+ antiporter. FEBS Lett. 365, 193–197. https://doi.org/10.1016/0014-5793(95)00455-i

Yu, J.-L., Grinius, L., Hooper, D.C., 2002. NorA Functions as a Multidrug Efflux Protein in both Cytoplasmic Membrane Vesicles and Reconstituted Proteoliposomes. J Bacteriol 184, 1370–1377. https://doi.org/10.1128/JB.184.5.1370-1377.2002

Žgur-Bertok, D., 2013. DNA Damage Repair and Bacterial Pathogens. PLoS Pathog 9. https://doi.org/10.1371/journal.ppat.1003711

Zhang, M., Ito, T., Li, S., Jin, J., Takeuchi, F., Lauderdale, T.-L.Y., Higashide, M., Hiramatsu, K., 2011. Identification of the third type of PVL phage in ST59 methicillin-resistant Staphylococcus aureus (MRSA) strains. FEMS Microbiol. Lett. 323, 20–28. https://doi.org/10.1111/j.1574-6968.2011.02355.x

Zhang, W., Luo, Y., Li, J., Lin, L., Ma, Y., Hu, C., Jin, S., Ran, L., Cui, S., 2011. Wide dissemination of multidrug-resistant Shigella isolates in China. J. Antimicrob. Chemother. 66, 2527–2535. https://doi.org/10.1093/jac/dkr341

Zhou, C., Bhinderwala, F., Lehman, M.K., Thomas, V.C., Chaudhari, S.S., Yamada, K.J., Foster, K.W., Powers, R., Kielian, T., Fey, P.D., 2019. Urease is an essential component of the acid response network of Staphylococcus aureus and is required for a persistent murine kidney infection. PLOS Pathogens 15, e1007538. https://doi.org/10.1371/journal.ppat.1007538

Zhu, Q., Kosoy, M., Dittmar, K., 2014. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. BMC Genomics 15, 717. https://doi.org/10.1186/1471-2164-15-717

Ziebuhr, W., Hennig, S., Eckart, M., Kränzler, H., Batzilla, C., Kozitskaya, S., 2006. Nosocomial infections by Staphylococcus epidermidis: how a commensal bacterium turns into a pathogen. Int. J. Antimicrob. Agents 28 Suppl 1, S14-20. https://doi.org/10.1016/j.ijantimicag.2006.05.012

9. Appendix

Accession No.	Isolate No.	Species	Area isolated
ERS2999996	1	*Staphylococcus haemolyticus*	East London Community
ERS2999997	27	*Staphylococcus sciuri*	East London Community
ERS2999998	33	*Staphylococcus sciuri*	East London Community
ERS2999999	59	*Staphylococcus sciuri*	East London Community
ERS3000000	74	*Staphylococcus sciuri*	East London Community
ERS3000001	75	*Staphylococcus sciuri*	East London Community
ERS3000002	93	*Staphylococcus haemolyticus*	East London Community
ERS3000003	99	*Staphylococcus haemolyticus*	East London Community
ERS3000004	105	*Staphylococcus haemolyticus*	East London Community
ERS3000005	109	*Staphylococcus sciuri*	East London Community
ERS3000006	207	*Staphylococcus hominis*	West London Community
ERS3000007	208	*Staphylococcus hominis*	West London Community
ERS3000008	209	*Staphylococcus hominis*	West London Community
ERS3000009	211	*Staphylococcus cohnii*	West London Community
ERS3000010	321	*Staphylococcus epidermidis*	East London Hospital
ERS3000011	327	*Staphylococcus epidermidis*	East London Hospital
ERS3000012	329	*Staphylococcus epidermidis*	East London Hospital
ERS3000013	343	*Staphylococcus cohnii*	East London Hospital
ERS3000014	349	*Staphylococcus cohnii*	East London Hospital
ERS3000015	355	*Staphylococcus epidermidis*	East London Hospital
ERS3000016	361	*Staphylococcus haemolyticus*	East London Hospital
ERS3000017	372	*Staphylococcus hominis*	East London Hospital
ERS3000018	373	*Staphylococcus haemolyticus*	East London Hospital
ERS3000019	385	*Staphylococcus hominis*	East London Hospital
ERS3000020	386	*Staphylococcus hominis*	East London Hospital
ERS3000021	387	*Staphylococcus hominis*	East London Hospital
ERS3000022	407	*Staphylococcus epidermidis*	East London Hospital
ERS3000023	435	*Staphylococcus epidermidis*	West London Hospital
ERS3000024	436	*Staphylococcus haemolyticus*	West London Hospital
ERS3000025	445	*Staphylococcus haemolyticus*	West London Hospital
ERS3000026	465	*Staphylococcus epidermidis*	West London Hospital
ERS3000027	475	*Staphylococcus epidermidis*	West London Hospital
ERS3000028	479	*Staphylococcus hominis*	West London Hospital
ERS3000029	492	*Staphylococcus haemolyticus*	West London Hospital
ERS3000030	506	*Staphylococcus haemolyticus*	West London Hospital
ERS3000031	538	*Staphylococcus haemolyticus*	West London Hospital
ERS3000032	620	*Staphylococcus hominis*	West London Hospital
ERS3000033	623	*Staphylococcus hominis*	West London Hospital
ERS3000034	631	*Staphylococcus epidermidis*	West London Hospital
ERS3000035	664	*Staphylococcus epidermidis*	West London Hospital
ERS3000036	673	*Staphylococcus epidermidis*	West London Hospital
ERS3000037	699	*Staphylococcus warneri*	West London Hospital
ERS3000038	700	*Staphylococcus warneri*	West London Hospital
ERS3000039	702	*Staphylococcus warneri*	West London Hospital
ERS3000040	711	*Staphylococcus epidermidis*	West London Hospital

ERS3000041	712	*Staphylococcus epidermidis*	West London Hospital
ERS3000042	713	*Staphylococcus epidermidis*	West London Hospital
ERS3000043	715	*Staphylococcus epidermidis*	West London Hospital
ERS3000044	716	*Staphylococcus epidermidis*	West London Hospital

Table 9.1: Summary of the Whole Genome Sequenced isolates