Whole-genome sequencing to aid diagnosis and clinical management of *Staphylococcus epidermidis* orthopaedic device-related infection

Kevin M. Cole

A Thesis Submitted in Partial Fulfilment of the Requirements of Brighton and Sussex Medical School for the Degree of Doctor of Philosophy

2022

Abstract

Device-related infection is a major complication of orthopaedic surgery and treatment can be complex and costly, requiring both antibiotic administration and surgery. Due to its ubiquity on the human skin *Staphylococcus epidermidis* is a common cause of orthopaedic device-related infection (ODRI). Confusingly, for the same reason, *S. epidermidis* is also a common contaminant of microbiological specimens and cultures. The culture of indistinguishable isolates from multiple specimens of peri-prosthetic tissue or fluid is a diagnostic indicator of ODRI. However, determining the relatedness of *S. epidermidis* isolates is difficult using conventional microbiological methods. Whole-genome sequencing provides a definitive measure of the relatedness of bacterial isolates and has the potential to be used to discriminate between *S. epidermidis* ODRI and contamination. Furthermore, the same assay allows detection of genetic elements associated with antimicrobial resistance and can be used to guide antibiotic therapy. This thesis aims to enhance our understanding of *S. epidermidis* diversity and provide the basis for a tool that can aid ODRI diagnosis and clinical management. This thesis is comprised of three chapters studying the genomics of *S. epidermidis* relating to population structure, the population dynamics of infection and antimicrobial resistance. The first study characterises the population structure of 192 colonies of *S. epidermidis* isolated from carriage sites of five healthy individuals. Multi-locus sequence typing was performed and compared with the genomic method of pairwise distance measurements. Pairwise distance measurements between colonies of different sequence types and from different individuals were used to define subtypes. The phylogeny of all isolates was established using ClonalFrameML and compared with sequence type and subtype data. A threshold of 100 single nucleotide variants (SNVs) was determined as a predictor of relatedness with a positive predictive value of 95%. Finally, the application of MLST, maximum-likelihood method, and subtyping were compared to characterise the population structure of the isolates.

The second study applied pairwise distance measurements to discriminate between 325 clinically defined infecting and contaminating isolates of *S. epidermidis* from peri-prosthetic specimen cultures. Haplotrees were then constructed to visualise the population structure of *S. epidermidis* associated with ODRI and specimen contamination. This study showed that population structures of *S. epidermidis* ODRI and *S. epidermidis* contaminating strains differ in ways that can be measured and visualised.

The third study genes and mutations associated with resistance to 13 antimicrobials were compiled into FASTA files. The FASTA files were used with the Basic Local Alignment Search Tool

to interrogate the genomes of 87 isolates of *S. epidermidis*.Results were compared with conventional antimicrobial susceptibility testing methods. Discordant results were further investigated with minimum inhibitory concentration testing, biochemical testing and interrogation of resistance genes. The genotyping methods yielded a sensitivity of 100% and specificity of 98% in correlating with the phenotypic methods.

These three studies demonstrate that whole-genome sequencing (WGS) can be used to enhance our understanding of bacterial populations in normal human carriage, infection and contamination. The methods used could be developed to provide a tool to aid diagnosis of *S. epidermidis* infections whilst also giving accurate information regarding the antimicrobial resistance of clinically important isolates.

Acknowledgements

I am greatly indebted to my supervisors Professors Martin Llewelyn, John Paul at Brighton and Sussex Medical School, and Daniel Wilson from the University of Oxford. Having you as my supervisors was my main prerequisite as I have always known you to be supportive, approachable and encouraging. Throughout the six long years your brilliant and positive supervision has meant I never once felt doubtful.

I would like to say thank you Dr. Bernadette Young from the University of Oxford whose advice and instruction on most of the analysis was invaluable. You have always been so generous with your time despite being so busy yourself. Thank you also for collecting very useful clinical data.

I would also like to thank Dr. Bridget Atkins working at the Nuffield Orthopaedic Centre for helping me to build a collection of bacterial isolates and your expertise on orthopaedic infection. I would like to thank all the laboratory staff at the John Radcliffe Hospital microbiology laboratory for diligently saving the cultures for me. Similarly, I would like to thank the amazing people at Royal Sussex County Hospital microbiology department and in the Nuffield Department of Medicine laboratory in Oxford for all their help and letting me use any equipment when needed.

Thank you to Dr. James Price at the Royal Sussex County Hospital for passing on your skills in analysis and for collecting clinical data. Similarly, thank you to Dr. Clare Gordon from the University of Oxford for passing on your expertise in resistance predictions methods.

Author Declaration

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to these or any other university for a degree and does not incorporate any material already submitted for a degree.

Signed: 

Date: 13/07/2022

Contents

List of Tables

Chapter 1 - Introduction

Chapter 2 – Materials and Methods

Chapter 3 - Measuring the diversity of *Staphylococcus epidermidis* in normal human carriage

Chapter 4 - Using whole-genome sequencing data to discriminate between *Staphylococcus epidermidis* orthopaedic device-related infection and contamination

Chapter 5 - Detecting genomic determinants of antimicrobial resistance in *Staphylococcus epidermidis*

List of Figures

Chapter 1 - Introduction

Chapter 2 – Materials and Methods

Chapter 3 - Measuring the diversity of *Staphylococcus epidermidis* in normal human carriage

Chapter 4 - Using whole-genome sequencing data to discriminate between *Staphylococcus epidermidis* orthopaedic device-related infection and contamination

Chapter 5 - Detecting genomic determinants of antimicrobial resistance in *Staphylococcus epidermidis*

Chapter 1
Introduction

1.1 *Staphylococcus epidermidis*

1.1.1 Taxonomic History

In 1882 Sir Alexander Ogston introduced the term "Staphylococcus" for grape-like clusters of bacteria associated with post-surgical infection (1). Friedrich Rosenbach adopted the name when in 1884 he described "*Staphylococcus pyogenes aureus"* and *"Staphylococcus pyogenes albus"* (2). Both forms, which differed in their golden or white colonial pigmentation, were considered by the early authors to be *bona fide* pathogens. In 1891, Welch reported the recovery of apparently non-pathogenic strains from the epidermis and from aseptic wounds, which he called "*Staphylococcus epidermidis albus*"(3). The married couple Charles and Anne Winslow, in 1908, applied the basonym (the earliest validly published name of the taxon) "*Albococcus epidermidis"* to:

> "A parasitic coccus living normally on the surface of the animal body. Occurs singly or in pairs, or irregular groups. Generally stains by Gram. Good to abundant, white surface growth. Moderate acid production in lactose, maltose, and glycerin media. Nitrates and neutral red reduced. Gelatin liquefied, somewhat slowly. Mannite not fermented"(4).

Alice Evans later transferred the species to the Staphylococcus genus to give the name *Staphylococcus epidermidis* (4). In 1940, Fairbrother recommended using coagulase production as a more reliable criterion than pigmentation for separating *Staphylococcus aureus* from other staphylococci (5) and in 1968 the lack of coagulase was used as a means of identifying the neotype (a new specimen designated as the type strain of a previously described species) strain American Type Culture Collection strain 14990 of *S. epidermidis* (4). However, until the 1970s, the names *"S. albus"* and *S. epidermidis* were applied generally to the coagulase-negative staphylococci (CoNS) (6), after which, important work by Karl Schleifer and Wesley Kloos characterized novel species, expanded the genus and more accurately delineated *S. epidermidis* from other taxa (7–9). More recently, morphological, physiological and biochemical arrays to describe and differentiate species of staphylococci have been succeeded by molecular methods which have helped better characterize the phylogeny and re-classify members to different genera, but *S. epidermidis* still remains an important member of the genus Staphylococcus.

1.1.2 Clinical Significance

S. epidermidis is ubiquitous amongst humans (10). It colonises the superficial mucosa and skin as a dominant part of the human normal flora (10).*S. epidermidis* can inhibit colonisation by the more pathogenic *S. aureus* by producing the serine protease Esp which can destroy *S. aureus* biofilms and inhibit its production (11). The introduction of Esp-producing *S. epidermidis* to the nasal cavity of volunteers carrying *S. aureus* resulted in the elimination of *S. aureus* colonisation (11). Although long regarded as primarily a commensal with little pathogenic potential, *S. epidermidis* and other CoNS are the microorganisms most commonly encountered in diagnostic microbiology laboratories (12). The ubiquity of *S. epidermidis* means that despite rigorous aseptic technique, a specimen can easily become contaminated by the normal flora of the clinical staff or the patient. Similarly, once at the microbiology laboratory, a specimen or culture may be accidentally contaminated during manipulation.

1.1.2.1 Emergence as a nosocomial pathogen

In 1958 Smith *et al.* found that between 1936 and 1955 5/338 (1.48%) cases of staphylococcal bacteraemia reported at their hospital were caused by "*S. albus*" and that approximately 1% of cases of subacute bacterial endocarditis (SBE) were caused by this organism (13). A few years later Brandt *et al.* drew attention to the increasing incidence of subacute bacterial endocarditis culturing "*S. albus*" (14). Increasingly common surgery involving the heart and great vessels was recognised as a factor associated with increasing numbers of staphylococcal SBE (14). By 1975 CoNS (including *S. epidermidis*) were associated with pyogenic, urinary-tract and prosthetic material infections (9). Kloos and Schleifer recognised their increasing importance and developed a simplified array to identify the ten staphylococcal species commonly isolated in routine medical microbiology (9). Freeman *et al.* (1987) observed a five-fold increase in CoNS-associated bacteraemia in a neonatal intensive care unit between 1975 and 1982 and felt this was not the result of an epidemic but an increase in the reporting of CoNS as a significant isolate (15). Between 1980 and 1989 the National Nosocomial Infections Study reported 70 – 297% increases in nosocomial bloodstream infection (BSI) mainly associated with *S. epidermidis* and other CoNS (16). In 2004 a survey of 24179 cases of nosocomial BSI implicated CoNS in the majority (31%) of cases (17). Multiple studies have found *S. epidermidis* to be the most common CoNS associated with BSI. (18–22). The increased range and use of implanted devices in medical practice has led to the increased incidence of CoNS and *S. epidermidis* infections. Inter-related to this is a change in patient demographics. The number of vulnerable hospital in-patients has increased both younger (e.g. premature neonates) and older, with more co-morbidities, chronic

illnesses, immunocompromisation, implanted medical devices and long-term hospital stays (23,24).

1.1.2.2 Infection types

Across many European Union (EU) and European Economic Area (EEA) countries CoNS are one of the leading causes of surgical-site infections and bloodstream infections (Table 1.1) (25). Within the CoNS, *S. epidermidis* is the most common organism associated with nosocomial and device-related infection (16,23) including bloodstream infection, prosthetic joint infection and continuous ambulatory peritoneal dialysis-related peritonitis (Table 1.2) Bloodstream infections are often catheter-related (26) with twice as many associated with central (CVC) as peripheral venous catheters (PVC) (17). Catheter-related bloodstream infections may arise at the percutaneous entry site with subsequent organism migration along the extraluminal catheter surface to the bloodstream (27,28). More rarely, they may result from infusate contamination or haematogenous seeding from other foci (27). *S. epidermidis*-associated BSI has been reported to complicate 4–5/1000 CVC insertions in the US (29). Short-term central venous catheters are at greater risk of infection per insertion days (27). Other risk factors include severity of underlying illness, granulocytopenia, integrity of the skin, presence of distant infection and catheter type (27). Clinical management usually includes catheter removal and antimicrobial therapy (27,28). *S. epidermidis* is the second most common cause of prosthetic heart valve infections accounting for around 13% of cases (the most common being *S. aureus*) (29). Prosthetic heart valve and pacemaker infections may result from accidental inoculation at the time of implantation or by haematogenous seeding (30). *S. epidermidis* accounts for 85% of CoNS-related native valve endocarditis which can have mortality rates similar to *S. aureus* (31). The next most common species is *Staphylococcus hominis* at 6% (31). CoNS are the most common cause of cerebrospinal shunt infections (48 – 67%) and of these *S. epidermidis* is the most common causal species (6). *S. epidermidis* is also a leading cause of neonatal bacteraemia where an immature immune system, the use of intravenous and umbilical catheters, and the presence of endemic hospital-adapted strains are risk factors (23,30,32–34). *S. epidermidis* and other CoNS have also been associated with infections of coronary stents, vascular grafts, intraocular lenses, breast implants, and continuous ambulatory peritoneal dialysis and urinary catheters (12,16,23,30,35). Other native tissue infections include osteomyelitis, otitis media, ocular infections and surgical-site infections (16,36–38). Moreover *S. epidermidis* is a major cause of orthopaedic device-related infections which will be discussed in detail below.

1.1.2.3 *S. epidermidis* Orthopaedic Device-Related Infection

Staphylococci are the most common organisms associated with prosthetic joint infection, accounting for over half of cases (25,39–42). Of these, *S. epidermidis* and *S. aureus* are the most common species (42–44). Whilst *S. epidermidis* and other CoNS produce less virulence factors than *S. aureus* they are still important pathogens associated with arthroplasty failure and high healthcare costs (40,45,46). Multidrug resistance and the ability to form biofilms can make *S. epidermidis* difficult to treat. Hellmark *et al.* (2009) found that of 33 strains isolated from PJI, 91% were resistant to more than three classes of antibiotics (47). Methicillin-resistant *S. epidermidis* is the most common coagulase-negative staphylococcus isolated from joint replacements (43,48)

	Coronary artery bypass graft	Hip prosthesis surgery	Knee prosthesis surgery	Laminectomy	ICU-acquired bloodstream infection*
Microorganism	n=269	n=1409	n=581	n=136	n=8659
Coagulase-negative staphylococci	26.4	18.9	17.6	15.4	20.5
Staphylococcus aureus	16.4	31.9	38.7	38.2	8.7
Klebsiella spp.	6.7	2.3	2.4	2.9	14.1
Enterococcus spp.	3.7	7.7	7.1	3.7	12.4
Pseudomonas aeruginosa	6.7	3.6	1.9	6.6	11.1
Escherichia coli	5.2	6.9	4.6	5.1	9.7
Enterobacter spp.	5.6	3	2.4	1.5	8.5
Candida spp.	2.2	0.5	0	2.2	8
Acinetobacter spp.	1.1	0.4	0	0	3.6
Serratia spp.	3.7	1.5	1.5	0.7	3.3
Other	22.3	23.3	23.8	23.7	0.1

Table 1.1. Percentage of microorganisms identified from coronary artery bypass graft, hip prosthesis surgery, knee prosthesis surgery and laminectomy surgical sites from 10 EU/EEA countries in 2017, and from ICU-acquired bloodstream infection from 10 EU/EEA countries in 2016. *Data from 626 isolates from Germany refer only to primary bloodstream infections (25,39).

Country	USA	USA	Denmark	Japan	Brazil	Greece	Spain	France	India	Italy	India	Germany	Turkey	India	S. Korea	USA	
Study	Gill *et al*, 1983 (49)	Kleeman *et al*, 1993 (50)	Jarlov *et al*, 1996 (18)	Kawamura *et al*, 1998 (51)	del Alamo *et al*, 1999 (19)	Petinaki *et al*, 2001 (52)	Cuevas *et al*, 2004 (53)	Sivadon *et al*, 2005 (54)	Singhal *et al*, 2006 (55)	Arciola *et al*, 2006 (56)	Chaudhury *et al*, 2006 (57)	Gatermann *et al*, 2007 (20)	Koksal *et al*, 2009 (21)	Jain *et al*, 2011 (22)	Shin *et al*, 2011 (58)	Roach *et al*, 2015 (59)	
Specimen	CS	CS	BSI	CS	BSI	CS	CS	BJI	CS	PJI(55)	CS	CRBI, BSI	BSI	CRBI	CAPD Peritonitis	CS	
No. of Isolates	**678**	**499**	**499**	**944**	**239**	**450**	**369**	**212**	**83**	**601**	**167**	**494**	**200**	**98**	**51**	**208**	**Mean %**
S. epidermidis	75	65	58	41	50	50	56	71	34	68	13	67	44	24	67	86	**54.3**
S. haemolyticus	7	13	9	15	10	15	5	2	13	7	72	12	12	37	12	11	**15.8**
S. hominis	5	7	12	5	12	11	18	4	2	8	6	9	9	3			**6.9**
S. capitis	2	4		5		4	3	6	10	5	1	1	8	12	4		**4.1**
S. warneri	3	4	2	3		4	4	7	1	5	1	2	4		8	2	**3.1**
S. lugdunensis		3		2		2		3	13	1	1	3	9				**2.3**
S. xylosus						4			6			1	5	13			**1.8**
S. caprae				14				2		1		1			6		**1.5**
S. saprophyticus	1	1	1	5		6	5						3			1	**1.4**
S. simulans	1	2		6		2	1	1			2	1	2	1			**1.2**
S. cohnii		1	1		6	1				3		1		8			**1.3**
S. schleiferi									12			1					**0.8**
S. pasteuri								2							2		**0.3**
S. auricularis													1				**0.1**
Not determined	6		17	4	22	1	8	2	9	2	4	1	3	2	1		**5.1**

Table 1.2. Percentage of CoNS species isolated from clinical samples 1983 – 2015 (Adapted from Piette *et al.* 2009 & Becker *et al.* 2014). CS = Clinical specimen, BSI = Bloodstream infection, BJI = Bone & joint infection, PJI = Prosthetic joint infection, CRBI = Catheter-related bloodstream infection, CAPD = Continuous ambulatory peritoneal dialysis.

1.1.2.4 Specimen contamination

Contamination of microbiological cultures can hinder successful diagnosis and treatment of infectious diseases. A contaminant may be mistaken as the cause of an aseptic disease resulting in inappropriate treatment. The presence of contaminants may also mask the presence of the true infecting pathogen which again could result in the administration of ineffectual antibiotics. The determination of whether a cultured organism is a contaminant or true cause of infection can be subjective yielded different diagnoses according to different laboratory staff and clinicians. Uncertainty of the significance of a cultured organism can also result in the unnecessary administration of antibiotics or, worse still, appropriate antibiotics not being administered to a true infection.

Measures, such as aseptic technique, the use of sterile disposable gloves and laminar flow cabinets are employed to prevent the contamination of liquid and agar cultures of diagnostic microbiological specimens. Surgically derived specimens from normally non-sterile sites may become contaminated with skin flora which can be mistaken for an infection by opportunistic pathogens. Sampling for prosthetic joint infection should be performed using a different set of sterile instruments for each specimen and transported to the laboratory in individual sterile containers (60,61). Whilst periprosthetic tissue is the standard specimen taken for intraoperative culture in the diagnosis of PJI the prosthesis itself would also be a useful specimen to culture, especially for the detection of a biofilm (62). However, handling the device can more often be associated with contamination than tissue samples which are smaller and more easy to manipulate in aseptic conditions (62).

While *S. epidermidis* and other CoNS may act as true pathogens, they are most frequently associated with contamination (63,64). Up to 82% of CoNS cultured from blood have been recorded as contaminants (64,65). Determining the clinical significance of CoNS isolated from normally sterile specimens is challenging and has important clinical consequences. Bates *et al.* (1991) found that, compared with negative blood culture results, false-positive results were associated with a median increase in hospital stay of 8 - 12.5 bed-days and median increased healthcare costs of over 33% (66). Souvenier *et al.* (1998) observed that in 59 cases of CoNS blood culture contaminations, 24 (41%) patients were unnecessarily administered antimicrobial therapy, with vancomycin being administered in 20 (34%) of the cases (64).

1.1.2.5 Determining the significance of *S. epidermidis* isolates

Given the potential for *S. epidermidis* to contaminate diagnostic samples, determining the clinical significance of isolates identified in the diagnostic laboratory requires careful correlation of laboratory and clinical data. For both bloodstream and prosthetic joint infection, diagnostic definitions have been proposed. For example, isolation of indistinguishable organisms from multiple cultures, time-to-positivity of culture and properties such as biofilm formation have been assessed as methods for laboratory-based diagnosis of bloodstream infection (63,64,67–75). For the diagnosis of prosthetic joint infection at hip or knee revision arthroplasty, Atkins *et al* (1998) demonstrated that detection of indistinguishable organisms from three or more operative samples was predictive of infection with a sensitivity of 65% and a specificity of 99.6% (61). More recently, the Musculoskeletal Infection Society (MSIS) published updated criteria for determining prosthetic joint infection based on clinical, biochemical, histological and microbiological observations in which two positive cultures of the same organism is considered a "major" factor in the determination of infection being present (76). Neither Atkins *et al.* in 1998 nor the MSIS in 2018 both regard the isolation of the same organism from multiple specimens at the site of the prosthesis as an indicator of PJI (61,76). However, neither expand on what is meant by "the same" or "indistinguishable". In practice, defining "indistinguishable" often relies on the methods to hand. Molecular methods can provide high resolution at determining relatedness of isolates; however, these are not always practical for use in routine diagnostic laboratories where determining relatedness is more often based on easily observable phenotypes which can be less accurate. Determining the relatedness of isolates of *S. epidermidis* is discussed in more depth in Section 1.1.5.2 Discrimination of *S. epidermidis* strains.

1.1.3 Virulence

S. epidermidis lacks the array of virulence factors harboured by *S. aureus*. The pathogenicity of *S. epidermidis* is usually associated with its ability to adhere to and colonise the polymeric materials of indwelling medical devices and then to protect itself from the host immune system and antimicrobial compounds (23,77). Most of these cellular factors are also associated with the commensal colonisation of the skin and external mucosa (Figure 1.1) and are not exclusively involved in human disease. Notwithstanding, several pathogenic toxins have been identified.

1.1.3.1 Toxins

Genomic analysis of the ATCC (American Type Culture Collection) 12228 reference strain found genes encoding δ-haemolysin and β-haemolysin to be the only toxin-producing genes present in this non-biofilm-forming-strain (78). Another *bona fide* virulence factor of *S. epidermidis* is the family of pro-inflammatory Phenol Soluble Modulins (PSM)s (79,80). PSMs are 22 to 44 amino acid peptides with a leucocyte cytolytic and cytokine-inducing ability (79,80). Although they have a role in the commensal lifestyle of staphylococci they are also associated with aggressive *S. aureus* infections and have been demonstrated to have a role in biofilm detachment and toxin mediated sepsis in *S. epidermidis* (78,81,82). Whilst most PSMs are encoded in the core genome, PSM-*mec* is encoded in the Staphylococcal Chromosome Cassette SCC*mec* which also harbours the methicillin-resistance conferring *mec* gene (83).

1.1.3.2 Biofilm formation

Biofilm-forming pathogens can cause devastating chronic infections in compromised individuals (84) Biofilm-enclosed bacteria are protected from antimicrobial therapy and the immune system in a way their planktonic counterparts are not (85). This can result in persistent or recurring disease requiring significant effort to treat or manage (85). Donlan and Costerton have defined biofilms as:

> "a microbially derived sessile community characterized by cells that are irreversibly attached to a substratum or interface or to each other, are embedded in a matrix of extracellular polymeric substances that they have produced, and exhibit an altered phenotype with respect to growth rate and gene transcription" (86).

During biofilm formation, initial attachment may involve non-specific, reversible binding to the device surface induced by van der Waals forces, electrostatic and hydrophobic interactions (23,30). The autolysin AtlE is thought to play an important role in this form of attachment with AtlE-deficient mutants showing impaired attachment (23,77). Initial attachment may also involve binding to the host matrix proteins fibrin, fibrinogen, fibronectin, vitronectin thrombospondin, collagen and von Willebrand factor that surround medical devices following implantation (23,87). This form of initial attachment is mediated by the microbial surface components recognizing adhesive matrix molecule (MSCRAMM) fibrinogen-binding protein Fbe (88,89). AtlE and the *S. epidermidis* cell surface protein SesC are also involved in this process (89). Attachment is followed by the accumulation into multi-layered cell clusters and is mediated by the Polysaccharide Intercellular Adhesin (PIA), known as poly-N-acetyl-

glucosamine (PNAG) in *S. aureus* (90). The role of PIA in the virulence of *S. epidermidis* is still unclear (90). PIA is a component of the exopolysaccharide matrix that surrounds the bacterial cell surface to protect against phagocytosis and killing by polymorphonuclear leucocytes (91). Deficient phagocytosis is thought to reduce the generation of an NF-kB mediated inflammatory response of the innate immune system (83)This matrix also retains nutrients and significantly reduces the penetration of vancomycin, oxacillin and cefotaxime (92). However, PIA production does not confer full impermeability to antibiotics which has been demonstrated by correlation between specific biofilm assays and standard disc diffusion methods, and observed penetration of fluorescent-tagged daptomycin (92). The accumulation-associated protein Aap was identified as mediating PIA-independent biofilm formation in clinical isolates of *S. epidermidis* (93). Aap prevents phagocytosis of PIA-independent biofilms (94) In *S. epidermidis* the Bap-homologous protein Bhp (also known as SesD) is a homologue of the biofilm-associated protein Bap in *S. aureus* and is thought to have a role in both attachment and accumulation of cells during biofilm formation but is not present in all strains (38,79,89,93). The mature biofilm is a multi-cellular agglomeration surrounded by a matrix of intercellular factors such as PIA, Aap, Bhp and extracellular (e)DNA (89). Similarities between the structure and metabolism of biofilms and tissues have been described in that different regions of biofilms exhibit different patterns of gene expression and contain channels which nutrients can circulate (84). Sections of mature biofilm may detach from the main body via a process mediated by the *agr* (accessory gene regulator) quorum-sensing system (29). In *S. aureus agr* down-regulates cell-surface proteins and up-regulates exoproteins (45). Detachment is also mediated by the production of PSMs which are also involved in the switch from quiescent to metabolically active modes (82). Biofilm detachment can result in transient bacteraemia and haematogenous seeding to other sites in the body including the heart.

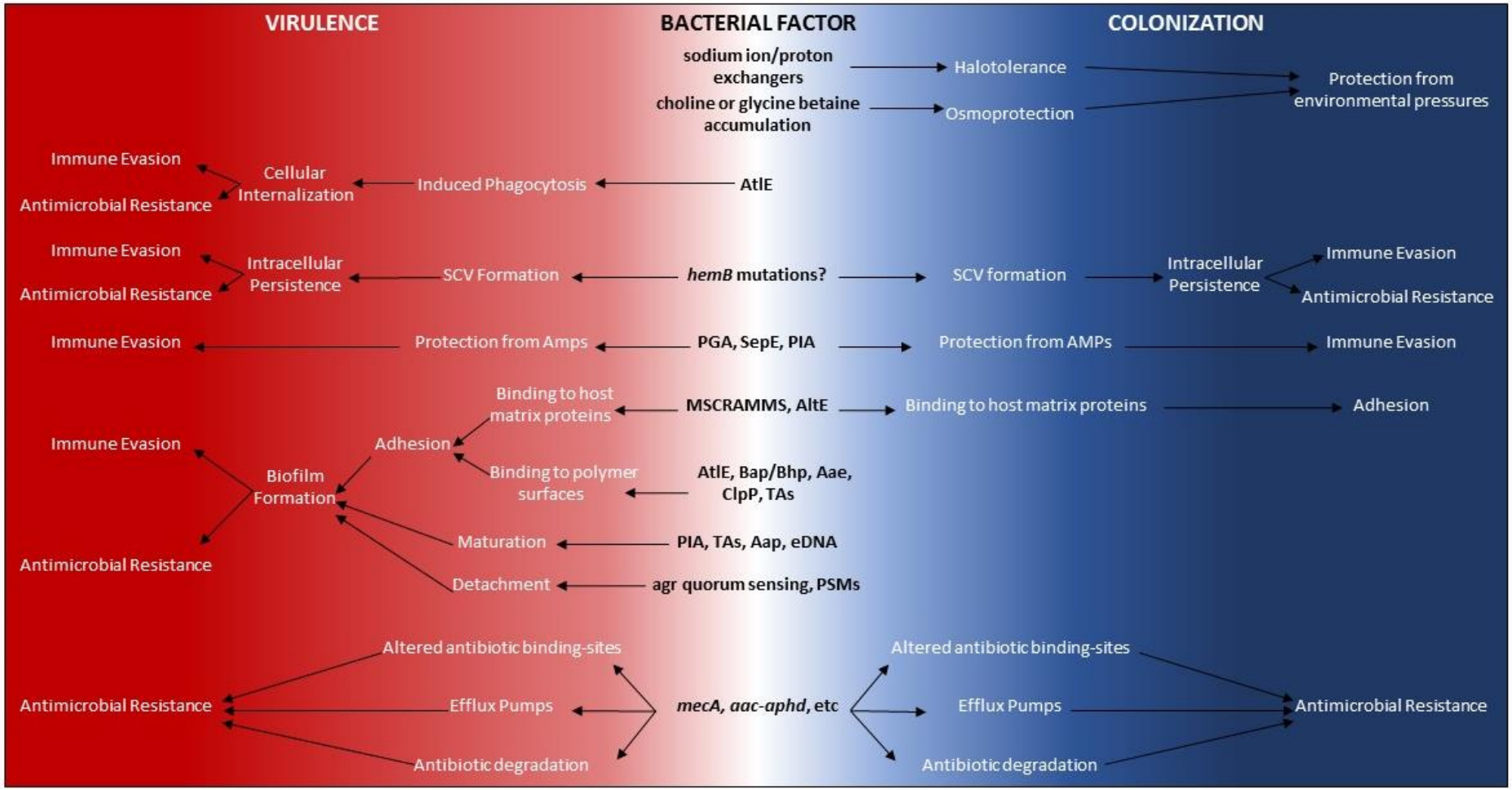


Figure 1.1. Bacterial factors associated with virulence and colonisation of *S. epidermidis*. Adapted from Otto. (2009) (95).

1.1.3.3 Genotypic determination of virulence

Many researchers have used gene-association studies to investigate determinants of nosocomial infection and hospital persistence (96–102). Table 1.3 summarises the main studies describing the distribution of virulence genes in nosocomial and commensal strains of *S. epidermidis*.

IS*256*

Insertion sequence IS*256* is thought to play a role in heterogeneous gene expression by reversible transposition into biofilm-associated genes (10,103) such as inducing reversible inactivation of the *icaADCB* operon (104). Multiple copies of IS*256* can be present in the *S. epidermidis* genome which has the ability to influence antimicrobial resistance gene expression (10,103). Gu *et al*. (2005) and Conlan *et al*. (2012) independently found IS*256* more commonly in *S. epidermidis* isolates of a clinical origin than commensal strains (99,103). However Rohde *et al*. (2004) found that IS*256* was found more commonly in commensal isolates from bone marrow transplant (BMT) patients than from invasive isolates from BMT patients suggesting the presence of IS*256* might be more indicative of nosocomiality than invasiveness (105).

icaADBC

The *icaADBC* operon encodes enzymes responsible for the formation of Polysaccharide Intracellular Adhesin (90). The production of PIA is has been identified as an important mechanism of biofilm accumulation following primary attachment to biosynthetic materials (90). However, clinically significant strains of *S. epidermidis* lacking PIA have been isolated and the accumulation-associated protein can act as a functional substitute (90). The *icaR* transcriptional regulator can repress transcription of the *icaADBC* operon in response to environmental conditions (104). A study by Frebourg and colleagues (2000) showed a higher presence of the *ica* operon in isolates from blood cultures and intravascular devices than in contaminating or colonising isolates (98). Moreover, again Gu *et al*. (2005) and Conlan *et al*. (2012) found *icaA* more commonly in *S. epidermidis* isolates of a clinical origin than commensal strains (99,103). However, the usefulness of the *icaA* gene as a marker of invasiveness has been challenged by Rohde *et al*. (2004) due to its higher presence amongst commensal isolates from BMT patients than clinical isolates from BMT patients (105).

bhp

The presence of the (Bap homologue protein) *bhp* gene has been associated with poor patient outcome of *S. epidermidis* orthopaedic device-related infection (100). The gene encodes a cell wall-associated biofilm protein that is thought to mediate PIA-independent intercellular

accumulation (100,106). It is homologous to the biofilm-associated protein (*bap*) gene in *S. aureus* (93). In *S. aureus* the *bap* gene encodes a 2276 amino acid cell-wall associated protein (107). All isolates of staphylococcus harbouring the *bap* gene are strong biofilm producers (77,107). Bap-positive isolates of *S. aureus* have an increased ability to colonise and persist in bovine mammary glands and, when in biofilms, are less susceptible to antibiotics (108) There is evidence of horizontal gene transfer, via the mobile pathogenicity island SaPIbov2, of the *bap* gene amongst *S. aureus* associated with chronic bovine mastitis (108). Biofilm-associated protein homologues have also been found in *Staphylococcus chromogenes, Staphylococcus xylosus, Staphylococcus simulans* and *Staphylococcus hyicus* as well as *Enterococcus faecalis*, *Burkholderia cepacia*, *Pseudomonas putida* and *Salmonella typhimurium* suggesting an extensive role in biofilm formation (77,108). In *S. epidermidis* he *bhp* gene has been found more commonly in nosocomial isolates than in healthy carriage isolates though without much difference between infecting and contaminating isolates (109).

aap

Expression of the 220 kDa accumulation-associated protein (Aap) as a truncated 140 kDa isoform is thought to the be the most important mechanism for PIA-independent biofilm formation in *S. epidermidis* (77,94). Expression of the full 220 kDa gene does not lead to PIA-independent biofilm formation suggesting the necessity of proteolytic cleavage (94). The *aap* gene mediates primary adhesion or biofilm accumulation or both and produces a different biofilm substructure than the *ica* operon encoded PIA (110). Two studies independently found the *aap* gene more frequently in carriage than in infecting isolates (109,111) and another study found the difference in frequency statistically insignificant (105), however the *aap* gene has also been associated with poor patient outcome of *S. epidermidis* ODRI (100).

agr

The staphylococcal accessory gene regulator locus (*agr*) is a quorum-sensing system that expresses cell-surface proteins for tissue colonisation during times of low cell density (88). When cell density increases, *agr* activation stops the production of surface proteins and induces the production of extracellular degradative exoenzymes (88). Autoregulatory feedback of the *agr* quorum-sensing system is mediated by a pheromone peptide when it binds to AgrC, a membrane-bound protein component of the system (112). The *agr* system comprises the genes *agrA, agrB, agrC* and *agrD* which are co-transcribed (RNA II) (111,112). In *S. epidermidis* three allelic groups of the *agr* system I, II and III which inhibit the activity of each other and the *agr*-systems of other staphylococcal species (111,112). *agr* type I is the type most frequently observed in clinical isolates of *S. epidermidis* (29,45,100) and currently non-typable forms have

been found implying the existence of more than the three characterised so far (45). The *agr* locus also houses the *hld* gene which encodes the short, heat-labile δ-haemolysin toxin that has lytic activity against different types of membrane (78).The marked prevalence of *agr* amongst both clinically significant and commensal strains means determination of type may be the most useful approach to determine association with infection however a study by Post *et al.* (2017) found no statistical significance between *agr* type clinical outcome of *S. epidermidis* ODRI unless the infection was acute in nature (100).

mecA

The methicillin resistance-conferring *mecA* gene is carried in the staphylococcal chromosome cassette SCC*mec* (113). SCC*mec* elements are diverse in structural organisation and genetic content and there are at least 14 known types to date in staphylococci (114–116). SCC*mec* consists of the *mec* gene complex and the cassette chromosome recombinase (*ccr*) gene complex (114,117). The *mec* gene complex contains *mecA*, and its regulators and the *ccr* gene complex encode recombinases of the invertase/resolvase family that mediate integration and excision of the cassette in and out of the chromosome at the *orfX* locus (113,114,117,118). The *mecA* gene encodes penicillin-binding protein PBP2a, a transpeptidase enzyme that has a lower affinity for ß-lactam antibiotics than PBP2 (119). Multiple studies have found the *mecA* gene more frequently in hospital isolated strains than community commensals (45,98,109,120). However, the use of *mecA* is a virulence marker is questionable since it has been detected similarly frequently in in-patient and specimen contaminating isolates as those deemed clinically significant (96,109).

ACME

The arginine catabolic mobile element (ACME) is a genetically diverse genomic island believed to have transferred from another staphylococcal species to *S. aureus* (121–124). Similar to SCC*mec*, the ACME integrate at the *orfX* locus of the staphylococcal chromosome (122,124). So far 5 types of ACME have been described in staphylococci all of which are present in *S. epidermidis* (124). Types vary according to the presence of *arc, kdp* and *opp3A* operons (122,124). The *arc*, *opp3* and *kdp* operons encode an arginine deaminase pathway, ABC transporter and potassium transporter, and an oligopeptide permease ABC transporter respectively (122,124). *In vivo* rabbit bacteraemia and mouse gut colonisation assays have demonstrated that ACME confers a survival advantage (122). ACME type I aids the colonisation of mucosal and skin sites of *S. aureus* strain USA300 and it is believed ACME confers a similar advantage in *S. epidermidis* (121,123). ACME is largely less frequently observed in clinical isolates than contaminating or commensal isolates (109,120,125). However, its ability to aid

survival and onwards transmission means this genetic element may play a role in the successful dissemination of other virulence factors in more virulent strains of *S. epidermidis*.

fdh

The formate dehydrogenase enzyme Fdh allows microorganisms to metabolise formate, a common substrate produced and consumed in various metabolic pathways in aerobic and anaerobic bacteria (126). Contrasting to genes associated with virulence, the formate-dehydrogenase (*fdh*) gene has been assessed as a genetic marker of commensalism due its higher prevalence in non-infecting isolates than clinical isolates (99,109,127,128). However its use as a marker of specimen contamination has been questioned by Tolo *et al.* (2016) who found it more frequently in—non-hospital carriage strains than infecting or contaminating strains (109).

Despite this wealth of data, none of the associations yet described between these various virulence determinants and disease are strong enough to inform decision making about the clinical significance of *S. epidermidis* isolated from individual patients (109). One recent genome-wide association study (GWAS) found virulence genes involved in adhesion and biofilm production were enriched in lineages associated with infection and that the isolates were better able to resist antibiotics and acidic pH values, adhere host matrix proteins such as collagen and survive in macrophages (129). However, a different GWAS could not associate virulence genes to PJI isolates with statistical significance but could detect genes and mutations conferring resistance to antimicrobials such as B-lactams, aminoglycosides and quinolones much more commonly in this group than contemporaneous nasal strains (130).

		Nosocomial									Community		
		Infection Associated			Contamination			Carriage			Carriage		
Genetic Element	Study	Tested (n)	Positive (n)	Positive (%)	Tested (n)	Positive (n)	Positive (%)	Tested (n)	Positive (n)	Positive (%)	Tested (n)	Positive (n)	Positive (%)
ica	Tolo *et al* (109), Cherifi *et al* (120), Hellmark *et al* (45), Gu *et al* (103), Rohde *et al* (96), Frebourg *et al* (98), Conlan *et al* (99), Yao *et al* (131), de Silva *et al* (33)	372	237	63.7	133	67	50.4	118	55	46.6	274	76	27.7
IS*256*	Tolo *et al* (109), Gu *et al* (103), Rohde *et al* (96), Conlan *et al* (99), Yao *et al* (131)	183	108	59.0	55	33	60.0	47	28	59.6	201	13	6.5
bhp	Tolo *et al* (109), Rohde *et al* (96),	75	27	36.0	55	21	38.2	25	4	16.0	55	9	16.4
aap	Tolo *et al* (109), Rohde *et al* (96), Yao *et al* (131)	158	119	75.3	55	36	65.5	25	23	92.0	99	70	70.7
mecA	Tolo *et al* (109), Cherifi *et al* (120), Rohde *et al* (96), Frebourg *et al* (98), Conlan *et al* (99)	237	200	84.4	94	65	69.1	25	24	96.0	175	30	17.1
ACME	Tolo *et al* (109), Cherifi *et al* (120), Granslo *et al* (125)	156	63	40.4	119	43	36.1				73	49	67.1
agr1	Hellmark *et al* (45)	61	35	57.4							24	10	41.7
agr2	Hellmark *et al* (45)	61	5	8.2							24	5	20.8
agr3	Hellmark *et al* (45)	61	21	34.4							24	8	33.3
*agr**	Hellmark *et al* (45)										24	1	4.2
fdh	Tolo *et al* (109), Conlan *et al* (99)	105	6	5.7	55	5	9.1				131	25	17.1

Table 1.3. Major studies assessing distribution of genetic markers with the potential to differentiate pathogenic, nosocomial and community carriage isolates of *S. epidermidis*. *Non-typeable *agr* (45).

1.1.4 Antimicrobial Resistance

Clinical isolates of *S. epidermidis* usually represent nosocomial strains which have adapted to the hospital environment. Nosocomial strains are usually resistant to multiple antimicrobials.

Methicillin

Of greatest clinical impact is resistance to methicillin, indicating resistance to essentially all beta-lactam antibiotics. Up to 80% of *S. epidermidis* isolates from device-related infections are methicillin resistant (24). Methicillin-resistant *S. epidermidis* (MRSE) strains are associated with hospital settings commensally colonising patients and staff and contaminating clinical specimens (96,109,132,133). In addition to *mecA, mecB, mecC* and *mecD* have all been reported in staphylococci (134–136). Moreover, *mec*-negative methicillin resistance has been reported in *S. epidermidis* (78).

Macrolides, lincosamides and streptogramin B

Macrolides, lincosamides and streptogramin B (MLS_B) are an important group of antibiotics for patients with a staphylococcal infection, especially those with allergy to beta-lactams. However MLS_B resistance is common amongst clinical isolates of *S. epidermidis* (137–139). MLS_B antibiotics target the bacterial 23S rRNA (140). Reduced affinity to 23S rRNA is mediated by plasmid-based *ermA, ermB* and *ermC* genes which encodes a methylase that causes the dimethylation of the adenine residue at position 2058 in 23S rRNA (140). The *msrA* gene encodes an ATP-dependant efflux pump that exports macrolides and streptogramin B (20,141). The macrolide phosphotransferase gene *mphC* is often found in association with the *msrA* gene (142,143). Benito *et al*. (2020) found the *mphC* gene more frequently in isolates of *S. epidermidis* from healthy individuals than from PJI or surgical field-related isolates (144). The healthy individual isolates represented a much greater number of STs suggesting that the plasmid-borne gene is widely disseminated (144). Less common is the *lnuA*-encoded lincosamide nucleotidyltransferase though the gene has still been found amongst nosocomial isolates (20,145,146).

Aminoglycosides

Aminoglycoside antibiotics bind to the 30S ribosomal subunit to inhibit protein synthesis with a bactericidal effect (147). Gentamicin and tobramycin are the most effective against staphylococci (147). Plasmid-borne *aac(60)-Ie-aph(200)-Ia, ant(40)-Ia,* and *aph(30)-IIIa* genes encode acetyltransferase (AAC), adenylyltransferase (ANT) and phosphotransferase (APH) enzymes in staphylococci (147,148) These enzymes alter aminoglycosides to reduce their ability to bind to the 30S ribosome subunit (141,147). Reported rates of aminoglycoside resistance

amongst clinical isolates of *S. epidermidis* have ranged from 56 – 71% (139,141,148). Rates are even higher if the isolate is also resistant to methicillin (139,148,149).

Tetracyclines

The tetracyclines (including tetracycline, doxycycline and minocycline) inhibit protein synthesis by blocking the association of aminoacyl- tRNA with the bacterial ribosome (150). The plasmid-borne *tetK* gene encodes an efflux protein which confers high-level resistance to tetracyclines (151). The gene has been found in clinical isolates of *S. epidermidis* at rates ranging from 15 – 98% (139,152,153). The *tetM* gene encodes a ribosomal protection protein to allow aminoacyl-tRNA associating with the acceptor site of the ribosome in the presence of tetracyclines (150,154). Soroush *et al.* (2016) detected the *tetM* gene in 5/80 (6.3%) of MRSE isolated from children (153).

Mupirocin

The topical antimicrobial mupirocin works by inhibiting the protein synthesis function of isoleucyl-tRNA synthetase (155). Resistance to mupirocin is mediated by two mechanisms in staphylococci. Non-synonymous point mutations in the isoleucyl-tRNA synthetase (*ileS*) gene causes low-level resistance to mupirocin, with an MIC ranging from 8 – 256 mg/L (156,157). Presence of the plasmid-borne *mupA* gene (previously known as *iles2*) confers a high-level resistance (MIC $\geq$512 mg/L) (155,158). The demonstration of *in vivo* transfer of *mupA* from *S. epidermidis* to *S. aureus* resulting in high-level resistance (159) and the identification of a chromosomally located *mupA* conferring a low-level resistance to mupirocin (156) suggest location and copy number of the gene can be clinically important.

Trimethoprim

The dihydrofolate reductase (Dhfr) enzyme is present in all organisms (160). Whilst the *dhfr* (*dfrB*) gene shows marked sequence divergence across species and evolutionary time (billions of years) structural homology of the enzyme remains high (228,229). Dhfr catalyses the reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate, an essential step in purine and thymidylate synthesis (160,161). Trimethoprim binds to and selectively inhibits bacterial Dhfr over mammalian Dhfr (160,162). Resistance to trimethoprim in staphylococci can be mediated by the acquisition of alleles *dfrA*, *dfrD*, *dfrG* or *dfrK* by horizontal gene transfer (163,164). Acquired *dhfr* gene variants confer high-level resistance to trimethoprim (163). The *dfrC* gene present in *S. epidermidis* is homologous to the *dfrA* gene in *S. aureus* with only 2 synonymous mutations separating them (165). Dual amino substitutions in the chromosomally transcribed Dhfr protein confers a higher level of resistance than single amino acid substitutions in *S. aureus*

(166,167). Resistance-associated *dfr* genes were found in 16/24 (75%) PJI specimens associated with staphylococci and other species by metagenomic sequencing (168).

Glycopeptides

The glycopeptides vancomycin and teicoplanin work by binding to the peptidoglycan precursors *N*-acetylglucosamine and *N*-acetylmuramic acid to prevent the transglycosylation process of peptidoglycan synthesis (169). High-level vancomycin resistance is associated with acquisition of the *vanA* gene (170). Expression of the *vanA* gene changes the dipeptide terminus from D-alanine-D-alanine to D-alanine-D-lactate to reduce the affinity of vancomycin a thousand fold (170). Transfer of the *vanA* gene from enterococci to staphylococci has been demonstrated as a feasible mode of acquisition and a public health concern (170,171). Glycopeptide resistance in staphylococci may also be in the form of heteroresistance (isogenic strains with different phenotypes observed amongst subpopulations) and/or reduced susceptibility (171,172). Exposure to vancomycin has been demonstrated to promote biofilm formation and cell wall thickening of *S. epidermidis* (173). Mutations in the RNA polymerase gene *rpoB* (also implicated in rifampicin resistance, see below), in particular H481Y and A621E, are associated with thickening of the *S. epidermidis* cell wall and reduced sensitivity to vancomycin (174). 35 point mutations identified at 31 loci were identified in a strain of *S. aureus* which showed *in vivo* evolution of vancomycin resistance with the MIC creeping from 1.0 mg/L to 8.0 mg/L) (175). In some strains resistance is inducible *in vitro* (171). The presence of biofilm has been demonstrated to reduce the penetration of vancomycin into *S. epidermidis* cells and therefore reduce its efficacy (92). *S. epidermidis* resistance or tolerance to glycopeptides is important as these antibiotics are commonly used as empirical therapy to serious CoNS infections (176). Resistance to vancomycin and teicoplanin has been found in strains isolated from BSI, ODRI and food (172,173,177–179).

Rifampicin

Rifampicin prevents DNA transcription by binding to the β-subunit of bacterial RNA polymerase encoded by the *rpoB* gene (180,181). Resistance to rifampicin is normally associated with point mutations in *rpoB* (47,181). Hellmark *et al.* (2009) found high-level rifampicin resistance was associated with one or two SNVs in the *rpoB* gene (47). Dual mutations D471E and I527M are thought to be the most common worldwide and also confer resistance to vancomycin (182). Mi *et al.* (2018) identified five different mutations in the *rpoB* gene associated with increased resistance to rifampicin but most came at a cost to fitness cost with mutants reverting back to wild type in the absence of rifampicin (181). The emergence of rifampicin resistance *in vivo* has been documented during a case of *S. epidermidis* related pacemaker infection demonstrating

the need to constantly monitor antimicrobial resistance during chronic infection (183). Rifampicin is thought to have a successful activity against staphylococcal biofilms (184,185) however it has also been shown to have reduced penetration of the polymeric matrix of *S. epidermidis* biofilms (186) and decreased activity against cells with low metabolic activity (187). Moreover surviving biofilm-located cells have been demonstrated to have a dramatically increased minimum bactericidal concentration (0.00025 to >128 mg/L) to rifampicin following exposure (186). Reports of rifampicin resistance in clinical isolates of *S. epidermidis* range from 4 to 58% (18,45,70,139,188,189).

Linezolid

Linzeolid belongs to the relatively new family of synthetic oxazolidinone antibiotics (190). Oxazolidinones inhibit protein synthesis by binding to the P site of the 50S ribsosomal subunit and preventing formation of the 70S subunit (190). This mode of action is different to that of tetracyclines, chloramphenicol, macrolides and lincosamides meaning staphylococci resistant to these antibiotics may still be susceptible to linezolid (190,191). Linezolid has been demonstrated to cause frameshifting and nonsense suppression during translation in *E. coli* suggesting this may have a role in its mode of action (192). Resistance to linezolid is still relatively rare amongst staphylococci and it is often used to treat multidrug resistant infections (189–191). Resistance to linezolid may result from modification of 23S rRNA part of the 50S subunit (189,190,193,194). A number of point mutations have been identified in 23S rRNA genes which confer resistance to linezolid in staphylococci (189,193,194). The plasmid-borne *cfr* gene produces a ribosomal RNA large subunit methyltransferase which methylates position 8 of adenine 2503 in 23S rRNA and confers resistance to linezolid, chloramphenicol, lincosamides and streptogramin$_A$ (189,194). Mutations in the *rplC*, *rplD* and *rplV* genes affecting the L3, L4, and L22 ribosomal proteins respectively have also conferred resistance to linezolid in staphylococci (189,194). Some strains of *S. epidermidis* harbour all three mechanisms (23S rRNA mutation, *cfr*-mediated and ribosomal L-protein mutation) of linezolid resistance (182,195). Linezolid resistance in *S. epidermidis* was first reported in 2004 in the US (196) and has since been reported in across the US, Europe and Brazil in a number of different sequence types (182,189,193,195,197,198)

Fluoroquinolones

Fluoroquinolone antibiotics interact with DNA gyrase and topoisomerase IV enzymes to stabilise DNA and type II topoisomerase complexes and thus inhibit DNA synthesis (199). DNA topoisomerase IV is more sensitive to fluoroquinolones in Gram-positive bacteria (200). Fluoroquinolones bind to the GyrA and GyrB subunits of DNA gyrase, and GrlA and GrlB

(previously called ParC and ParE respectively) of DNA topoisomerase IV (201,202). Mutations in the *gyrA, gyrB, grlA* and *grlB* genes modify the fluoroquinolone binding sites and prevent activity (201,203). Mutations in the *gyrA* and *grlA* genes are more common than in *gyrB* and *grlB* (201). Mutations present in both the *gyrA* and *grlB* confer high-level resistance to ciprofloxacin in *S. aureus* (203). Combinations of mutations in the *gyrA, gyrB, grlA* and *grlB* genes are known to have varying effects on fluoroquinolone resistance, however the relationship between genotypic and phenotypic relationships are still not fully understood (204). Similarly, the NorA efflux pump is associated with resistance to fluoroquinolones although its effect is still not fully known (201,204). Analogous modes of resistance have been found in clinical isolates of *S. epidermidis* to that of *S. aureus* (202,205,206). A survey of over 80,000 *S. epidermidis* BSI isolates in the US showed that ciprofloxacin resistance increased steadily from 58.3 to 68.4% in the US between 1999 and 2012, whilst levofloxacin resistance increased from 57.1 to 68.1% (196). Hellmark *et al.* (2009) determined that 26/33 (79%) of PJI isolates had resistance or intermediate susceptibility to norfloxacin, ciprofloxacin and moxifloxacin (47) Excretion of ciprofloxacin in sweat has been suggested as a possible cause of the emergence of resistance in *S. epidermidis* on the human skin which may then go on to cause nosocomial infection (207).

Fusidic Acid

Fusidic Acid inhibits bacterial protein synthesis by binding to translocase elongation factor G (EF-G) and preventing its release from the ribosome (208,209). In staphylococci EF-G is encoded by the *fusA* gene (210). High-level resistance (HL-R) ($\geq$256 mg/L) in *S. epidermidis* is associated with the point mutation L461K in the *fusA* gene of (210). Other *fusA* point mutations have been associated with HL-R in *S. aureus* (211). The FusB family of proteins bind to EF-G on the ribosome to allow dissociation (212). FusB proteins are encoded by *fusB* and *fusC* which can be transferred horizontally between staphylococci by plasmid and phage, and SCC cassette respectively (212). The *fusB* and f*usC* genes both confer a lower-level fusidic acid resistance than the *fusA* point mutation L461K in *S. epidermidis* (210). Mutations in the *fusE* (*rplF*) encoding ribosomal protein L6 confers fusidic acid resistance in *S. aureus* and is associated with haemin auxotrophy and a small-colony variant phenotype (213). Resistance rates ranging from 0 to 32% have been documented across Europe and in Japan (214).

1.1.5 Laboratory Identification

1.1.5.1 Discrimination of *S. epidermidis* and other Coagulase-negative Staphylococci

Members of the *S. aureus* complex are more pathogenic than other members of the genus and are the only ones to produce coagulase, staphylococcal protein A and DNase which serve as important biochemical markers. After this differentiation, identification of *S. epidermidis* is largely based on discriminating between the phenotypic features of it and other coagulase-negative staphylococci, a relatively easy process nowadays. Commercial platforms using Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry (MALDI-TOF MS) such as the MALDI Biotyper system (Bruker Daltonik GmbH, Leipzig, Germany) or the VITEK MS (bioMérieux, Marc' l'Etoile, France) provide relatively cheap, quick and accurate tools for species identification (215–217). The criteria of a Biotyper score of 2.0 – 2.299 for implying a secure genus and probable species identification has been demonstrated to give 100% frequency in accurately identifying *S. epidermidis* (and *S. aureus*, and non-*S. epidermidis* CoNS) however scores of >2.3, implying highly probable species identification are achieved much less frequently (218).

16S rRNA gene sequencing has been successfully employed to identify staphylococci and investigate the phylogeny of the genus (219–222). Identification methods based on RNA polymerase B (*rpoB*), superoxide dismutase A (*sodA*) and elongation factor Tu (*tuf*) gene sequencing analysis have been proposed (223–227) but have not replaced 16S rRNA gene sequencing (228). Genotypic methods are reliable, but they have not yet been applied in routine diagnostic settings.

1.1.5.2 Discrimination of *S. epidermidis* strains

Distinguishing between strains of various pathogenic bacteria in clinical laboratories often relies on phenotyping methods (5,7,9,229). Morphological differences (such as colony size, colour or haemolytic ability) between strains may be observed on agar cultures. Antibiotyping uses the profile of susceptibility to the antibiotics tested as routine practice. Both methods can potentially provide useful information at no or low additional extra cost. However, morphological features and antibiotypes may be shared by genetically unrelated strains meaning they may be accidently regarded as monoclonal. Conversely, genes that determine antimicrobial resistance or morphology may vary in their expression and obscure epidemiological links between related isolates (230). To avoid such pitfalls, guidance documents

have been published regarding the validation and application of epidemiological typing systems (231,232). Moreover, genotypic methods have been established to provide more accurate information on the relatedness of strains.

Multi-locus sequence typing (MLST) for *S. epidermidis* exploits the sequence variation present in specific loci in seven house-keeping genes (listed in Table 1.4.) (233). Different alleles are assigned a number so that a Sequence Type (ST) based on the numerical profile of the seven loci can be assigned (233). Therefore a sequence type is a strain (or clone) defined by the alleles at the seven loci (234). Sequence types can be grouped according to the number of shared alleles into Clonal Complexes (CC)s. Clonal complexes are usually comprised of a single predominant ST and a number of less common STs differing by a single allele or more (234). The PubMLST.org database is the internationally accepted database for allele sequences and sequence types (235). Whilst MLST has proved invaluable in the characterisation of the population structure Miragaia *et al.* (2008) determined that it was not the optimal method for defining *S. epidermidis* clones (236). They found the technique was not sufficiently discriminatory to detect the short-term evolutionary changes that can define clones and that, surprisingly, a combination of pulsed-field gel electrophoresis (PFGE) followed by SCC*mec* typing could provide both reliable information on short-term epidemiology and predict long-term evolutionary history (236). The authors speculate that because the rate of SCC*mec* acquisition is lower than the rate of events generating new PFGE subtypes the specific SCC*mec* type is associated, if transiently associated with a characteristic PFGE subtype generating greater concordance than MLST (236).

Pulsed field gel electrophoresis employs restriction endonuclease enzymes to cleave chromosomal DNA into fragments which are then subjected to gel electrophoresis to separate them (237). The fragment size profiles are strain-specific and so can be used as a discriminatory typing method (237). PFGE has been determined to be a more suitable technique than MLST for discriminating local nosocomial clones (238). Studies have shown that, unlike MLST, clinical isolates can have a broad range of PFGE pulsotypes (117,120,239). Although it is highly discriminatory, PFGE alone may not detect non-clonality between isolates (240) which could lead to inaccurate diagnosis of infection. Polyclonality of *S. epidermidis* from bloodstream and prosthetic joint infections have been indicated by PFGE (241,242). However, such results may reflect heterogeneity as a result of plasmid loss (243) chromosomal rearrangements (241,244) or changes in the insertion sequence IS*256*-specific hybridization patterns (244). Genotypic variability may be observed via differing PFGE patterns (241,245) or plasmid profiles. When

combined with staphylococcal chromosome cassette SCC*mec* cassette typing it provides results similar to those obtained using MLST (236).

Gene	Alleles (n as of 21st June 2021)	Length (bp)	Protein	Molecular Function/s
arcC	91	465 – 466	Carbamate kinase	Carbamate kinase activity in carbamoyl phosphate degradation
aroE	95	420	Shikimate dehydrogenase (NADP(+))	NADP binding, Shikimate 3-dehydrogenase (NADP+) activity
gtr	92	435 – 438	ABC transporter	ATPase-coupled transmembrane transporter activity
mutS	63	412	DNA mismatch repair protein MutS	ATP binding, damaged DNA binding, mismatched DNA binding in DNA repair
pyrR	72	428	Bifunctional protein PyrR	Guanine phosphoribosyltransferase activity, guanine phosphoribosyltransferase activity, hypoxanthine phosphoribosyltransferase activity, RNA binding, uracil phosphoribosyltransferase activity
tpiA	77	219 – 424	Triosephosphate isomerase	Triose-phosphate isomerase activity
yqiL	80	416	Acetyl coenzyme A acetyltransferase	Transferase activity, transferring acyl groups other than amino-acyl groups

Table 1.4. Description of loci employed in the current MLST scheme for *S. epidermidis* (246–254)

1.1.6 Population Structure and Epidemiology

The population structure of *S. epidermidis* can be characterised by various means. The population structure in human carriage may reflect how different strains sharing the same ecological niche interact and how new strains emerge from colonising populations. Establishing genetic relatedness of globally reported strains of *S. epidermidis* can show the true diversity of the organism, its phylogeny and how clonal lineages are represented. Investigating clinically significant isolates has the potential to reveal genetic markers of pathogenicity or nosocomial adaptation. Clinical isolates of *S. epidermidis* have been found to generally belong to a number of successful hospital-adapted clones that have spread within and between countries (10,133,182,238).

1.1.6.1 Population structure of clinical isolates according to multi-locus sequence typing

The current MLST scheme for *S. epidermidis* was devised by Thomas *et al.* (2007) (246) and successfully applied to characterise 217 nosocomial isolates (238). From the 217 isolates 74 different STs were identified although 74% fell into a single clonal complex (CC2) comprised of 39 sequence types (STs) of which ST2 was the most common (67/217) (238). At the time it was inferred that ST2 was the founding sequence type (238) however Rolo *et al.* (2012) later concluded that ST5 was the founder (117). The predominant nosocomial strains ST2, ST23 and ST27 more frequently harbour virulence-associated genetic elements both within clinical and hospital-carriage isolates (111,238,255,256). Thomas *et al.* (2014) applied Bayesian clustering of the available *S. epidermidis* MLST data to assign all STs into 6 genetic clusters (GCs) (128). Isolates were then assigned to one of the GCs which were strongly associated with either a nosocomial, commensal or generalist origin (128). Tolo *et al.* (2016) then applied the genetic clustering to predict isolate source, with an accuracy of 80% for predicting nosocomial origin but only 45% for predicting infection or contaminant origin (109). At the time of writing there are 1124 STs of *S. epidermidis* illustrating how the known population structure has expanded dramatically since the introduction of the scheme (254). MLST has been useful in identifying globally distributed successful lineages. Lee *et al.* (2018) identified two lineages of ST2 and one of ST23 that have become highly successful nosocomially and have acquired important drug-resistant mutations and spread globally (182). These lineages have been isolated from healthcare institutions in 24 countries in Europe, the United States and Australia between 2000 and 2018 with many clones resistant to rifampicin and heteroresistant to glycopeptides (182). A study by different authors identified nine isolates (eight from infection site and one

from the nose) of one of the drug resistant ST2 lineages described by Lee and colleagues (2018) (182,257) suggesting these lineages may be more widespread than originally thought.

1.1.6.2 Epidemiology according to pulsed-field gel electrophoresis

The use of PFGE in assessing the epidemiological significance of *S. epidermidis* was first reported in 1992 (258,259). Since then it has been has been widely adopted for hospital outbreak investigation or comparison of isolates within culture sets (67,119,239,240). In 2010 PFGE was used to determine an outbreak of two major clone types of *S. epidermidis* in hospitals in two different cities in the United States (198). Type one was isolated from both hospitals and type two was isolated from one hospital (198). The *cfr* gene and an insertion in the L4 ribosomal gene were detected amongst isolates of both types and the G2576T mutation was detected amongst isolates from type two, these genetic determinants are associated with resistance to linezolid, which had been given to all affected patients within six months (198). In 2016 PFGE and MLST were used together to characterize the diversity of *S. epidermidis* from BSI in hospitals across Belgium (139). The authors found eight different pulsotypes comprising 37 isolates that belonged to CC2 were isolated from more than one hospital (139). This demonstrated how MLST could be used to determine an epidemic clonal lineage and PFGE could then further discriminate between strains within the lineage (139).

1.1.6.3 The population structure of commensal carriage

A great deal of work investigating the small number of strains of nosocomial strains reflect their importance and the need to understand them to treat and prevent disease. However, a larger number of strains of *S. epidermidis* are commensal to humans. and their ubiquity on the epidermis means they are potential contaminants of clinical specimens. Therefore, an understanding of these can have a use in the diagnosis of infection. Whilst there is a paucity of research using methods like MLST and PFGE to characterise commensal carriage of *S. epidermidis* the recent interest in the human microbiome has yielded work using more advanced methods.

Commensal carriage of *S. epidermidis* is characterised by diversity both within and between people (99,260,261). Metagenomic analyses have shown intra-body distribution to be a more important driver of variation than inter-body distribution with different subtypes preferring different microenvironments on the skin (260). Heterogeneous communities at different niches are relatively stable over time with little acquisition of external strains (261). The

marked internally derived variation maybe a result of the open pan-genome of *S. epidermidis*. Pan-genome sequence analysis has shown up to 20% of genes are variable between strains (99,262). This variable genome is enriched with genes involved in replication, recombination and repair including mobile genetic elements harbouring recombinase and integrase genes (99).

1.2 Orthopaedic Device-related Infections

Implanted orthopaedic devices include prosthetic joints of the hip and knee, and less commonly of the ankle, shoulder or elbow. They also include devices used in the surgical treatment of trauma, including the fixation of fractured bones. Devices may be comprised of metal, ceramic or polymeric materials (263). Great efforts are taken to prevent infection of implanted devices such as ultra-clean-air theatres, peri-operative antibiotics and use of cements loaded with antibiotics (264) as well as the surgical site and surgeon decontamination and the wearing of surgical clothing. However, despite best efforts infections do occur which can be very complicated and expensive to treat. One 15 year prospective study in the United Kingdom identified deep infections developing in 34/5947 (0.57%) primary hip replacements and 41/4788 (0.86%) knee replacements (264). Infection rates can be higher following fracture fixation surgery with studies reporting 1 – 2% for closed fractures and up to 30% for open fractures (265).

Prosthetic-joint infections can be classified according to their temporal relationship with prosthesis insertion and the presumed mechanism of device infection. Early (<3 months post-surgery) and delayed (3 – 24 months post-surgery) infections are usually associated with direct inoculation at the time of device insertion (264,266). Late (>24 months post-surgery) infections more often result from haematogenous seeding (264,266). Indwelling medical devices such as orthopaedic prostheses provide a nidus for bacterial growth and lack a vascular system that can deliver an immune response or sufficient antibiotics, making them vulnerable to infection by opportunistic pathogens (35).

Treatment of ODRI always requires surgical intervention since antibiotic treatment alone is not able to eradicate biofilms (267). Sessile, biofilm producing organisms are aggregated in an exopolysaccharide matrix which large molecules may not be able to penetrate (86,92). Moreover, cells within the biofilm can have a lower metabolic activity which can reduce the interaction between the antimicrobial compound and its target (82). Both of these factors can lead to a persistence of the organism in the presence of the host immune response or antibiotics (77,86,268,269). Surgical options for infections following hip or knee arthroplasty

include washout of the area, “Debridement, Antibiotics and Repair” (DAIR), one or two-stage revisions or even amputation (270,271) all of which can be complex and costly. Treatment of ODRI is successful in 57 – 88% of cases (267).

1.2.1 Diagnosis

Diagnosis of prosthetic joint infection can be challenging. Atkins *et al*. (1998) demonstrated that detection of indistinguishable organisms from three or more operative specimens taken with a different set of surgical instruments was 96.4% predictive of infection (61). Using the cut-off of two or more indistinguishable organisms decreased the probability of infection to 80.6% and this result was largely due to patients with three or more culture-positive specimens being included in this group as patients that had exactly two indistinguishable cultured organisms yielded a probability of infection of only 25.2% (61). In 2018, the Musculoskeletal Infection Society published updated criteria for determining prosthetic joint infection based on clinical, haematological, histological and microbiological observations (Table 1.5), which includes a scoring-based system and a suggestion to consider molecular diagnostic testing. Application of this system yielded a sensitivity of 97.7% and specificity of 99.5% in differentiating PJI from aseptic revisions during external validation (76). Interestingly, the 2021 version of the UK Standards for Microbiology Investigations (SMI): Investigation of orthopaedic implant associated infections by Public Health England does not reference the MSIS criteria but does reference the Atkins *et al.* study (60). The SMI discusses the challenges in defining the term “indistinguishable” as indicated by similarities in an extended antibiogram, no doubt reflecting common practice in UK diagnostic laboratories where genotypic typing methods such as MLST and PFGE are not commonplace (60).

Major criteria (at least on of the following)	Decision
Two positive cultures of the same organism	Infected
Sinus tract with evidence of communication or visualisation of the prosthesis	

	Minor criteria		Score	Decision
Preoperative Diagnosis	Serum	Elevated CRP or D-Dimer	2	≥6 Infected
		Elevated ESR	1	
	Synovial	Elevated synovial WBC or LE	3	2-5 Possibly Infected
		Positive Alpha-defensin	3	
		Elevated Synovial PMN %	2	0-1 Not Infected
		Elevated synovial CRP	1	

	*Inconclusive pre-op score or dry tap	Score	Decision
Intraoperative Diagnosis	Preoperative Score	-	≥6 Infected
	Positive Histology	3	4-5 Inconclusive**
	Positive Purulence	3	≤3 Not Infected
	Single Positive Culture	2	

Table 1.5. New scoring based definition for periprosthetic joint infection. MSIS (2018).

Proceed with caution in: Adverse local tissue reaction, crystal deposition disease, slow growing organisms. *For patients with inconclusive minor criteria, operative criteria can also be used to fulfill definition for PJI. **Consider further molecular diagnostics such as next-generation sequencing. ESR: Erythrocyte Sedimentation Rate. WBC: White Blood Count. LE: Leukocyte Esterase. PMN: Polymorphonuclear Neutrophil. CRP: C-Reactive Protein. (76).

Infection after fracture fixation (IAFF) is different to prosthetic joint infection and therefore diagnosis has different requirements (Table 1.6) (267). However, like the criteria listed in Table 1.5. proposed by Parvizi *et al*. (2018) diagnosis of IAFF can be made using clinical, histological and haematological findings and the culture of indistinguishable organisms from more than one specimen from the site (267).

Confirmatory criteria include;
• Purulent drainage from the wound or presence of pus during surgery • Phenotypically indistinguishable pathogens identified by culture from at least two separate deep tissue/implant (including sonication-fluid) specimens taken during an operative intervention. • Presence of pathogens in deep tissue, taken during an operative intervention, as confirmed by histopathological examination.

Suggestive criteria include;
• A pathogenic organism identified by culture from a single deep tissue/implant specimen taken during an operative intervention • Elevated serum inflammatory markers: especially suggestive in a case of secondary rise (after an initial decrease), or a consistent elevation over a period of time (ESR, WBC count, CRP)

Table 1.6. Criteria proposed by Metsemakers *et al*. (2018) for the diagnosis of fracture-related infection (265,272).

1.2.1.1 Sampling and Culture

The positive and negative predictive value of culture in diagnosis of ODRI is strongly reliant on sampling factors. Prior to sampling, antibiotics should be stopped, ideally two weeks beforehand to reduce the chance of false culture-negative results (61,273). Tissue, synovial fluid and swabs are common specimens taken at the time of debridement or device explantation (61,273,274). Drainage from the wound site or sinus tract is not recommended as a specimen type due to high rates of contamination (274). Sonication and culture of the explanted device to remove biofilms has been assessed with varying degrees of success, however due the size of some devices the procedure requires more specialist equipment, can be cumbersome and associated with contamination (62,275). Implanted devices are more commonly infected with organisms associated with the normal flora of the human skin which can also be common contaminants of specimens and cultures. Therefore, sampling must be done aseptically, using a sterile set of surgical instruments for each specimen and placed in separate sterile containers (61). Specimens should then be sent to the laboratory for immediate culture to reduce the chance of fastidious organisms dying. Specimens should be cultured onto nutritionally rich media such as blood agar and chocolate agar and incubated in CO_2 alongside anaerobic culture on a suitable media. In addition, inoculation into a

nutritionally rich broth and incubation prior to sub-culture is recommended (60). Alternatively, specimens may be directly inoculated into aerobic and anaerobic blood culture bottles for the automated detection of growth (60,276).Enriched broth and blood culture media are useful in recovering sessile or low numbers of bacterial cells, however it can be harder to discriminate between infecting and contaminating organisms than a direct agar culture where contaminants can be detected easier by the trained microbiologist. Culture should be prolonged to allow the growth of slow-growing organisms (277).

1.2.1.2 Identification

As mentioned in section 1.1.5 Laboratory Identification, identifying cultured organisms to species level has become relatively simple with the advent of mass-spectrometry technology. Cultured organisms then undergo antimicrobial susceptibility testing to both determine which antibiotics may be suitable for treatment and determine antibiogram which can be used as surrogate marker for strain relatedness as previously discussed in section 1.1.5.2 Discrimination of *S. epidermidis* strains. Other typing methods such as MLST and PFGE to determine the relatedness of bacterial strains has also been discussed previously.

1.3 Hypothesis and Aims

I hypothesis that applying modern analytical methods to whole-genome sequencing data of *S. epidermidis* will expose novel information about the population structure of the organism in healthy carriage, infection and contamination.

Colonising and infecting populations of *S. epidermidis* are not strictly monoclonal as diversity increases over time. Diversity is measurable as the accumulation of single nucleotide polymorphisms in the genome. Using whole-genome sequencing to measure diversity within and between colonising populations and strains will help define these groups. Moreover, measuring diversity of infecting and contaminating populations may highlight differences which could be exploited to discriminate true *S. epidermidis* infection and specimen or culture contamination. The same whole-genome sequencing data can also be used to detect genetic determinants of antimicrobial resistance to supplement or supplant phenotypic testing.

Overall, the aim of this thesis was to characterise the diversity of *S. epidermidis* in normal human commensalism, specimen contamination and device-related infection so that a method could be developed which could potentially aid clinical diagnosis of infection, and then develop a method in which antimicrobial resistance could be detected. It was therefore intended to further advance an understanding of the population structure of *S. epidermidis* and produce tools which could have a useful clinical impact on the detection and management of this emerging infection.

Aim 1 - Measuring the diversity of *S. epidermidis* in normal human carriage

This body of work was intended to characterise the diversity of *S. epidermidis* that exists within and between populations colonising healthy individuals and the diversity that exists within and between strains defined by multi-locus sequence typing. It was intended to determine if measurable thresholds exist which can differentiate *S. epidermidis* subpopulations and strains.

Aim 2 - Using whole-genome sequencing data to discriminate between *S. epidermidis* orthopaedic device-related infection and contamination

The aim of this work was to measure the diversity of populations of *S. epidermidis* from rigorously defined cases of orthopaedic device-related infection and incidental specimen

contamination. These measurements were then going to be analysed to determine if any differences could be used to discriminate between the two sample groups.

Aim 3 - Detecting genomic determinants of antimicrobial resistance in *S. epidermidis*

This pilot study was intended to determine if the use of whole-genome sequencing to detect known genetic determinants of antibiotic resistance in *S. epidermidis* correlates with the currently adopted methods of phenotypic antimicrobial susceptibility testing.

Chapter 2
Materials and Methods

2.1 Culture and Identification

2.1.1 Bacterial Culture

To investigate the phenotypic and genotypic diversity of populations isolated from original specimens, single colonies similar in appearance were sub-cultured onto a Columbia Agar & Horse Blood (CBA) plate (Oxoid Ltd, Basingstoke, UK) and incubated aerobically overnight at 35°C

2.1.2 Identification of *S. epidermidis*

Colonies of *S. epidermidis* were initially screened according to their visual appearance; circular, raised, white, non-haemolytic colonies with a sharp border and generally growing to 1 – 2mm following overnight incubation on CBA or Columbia Cap Selective Agar. Suspect colonies underwent identification using matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-ToF MS), (Bruker Daltonik GmbH, Leipzig, Germany). In summary, a small amount of the colony is smeared onto the well of a target plate. A 0.5µl drop of matrix is then applied and allowed to air-dry. The matrix is α-Cyano-4-hydroxycinnamic acid (HCCA) which crystallises to enable measurement of peptides and proteins from 0.7 to 20kDa (278). The target plate is inserted into the instrument where laser ionisation/desorption charges the analytes which are then accelerated through a vacuumed flight tube. The time-of-flight (ToF) of peptides and/or proteins through the flight tube is used generate a peptide-mass fingerprint (PMF). This PMF (or spectrum) is then compared with the Bruker Biotyper database of known organism proteomes to give an identification with a corresponding quality score. Colonies that identified as *S. epidermidis* with a score of 2.0 or greater were considered suitable for initial analysis. Isolates undergoing further investigations were identified using WGS data

2.2 Antimicrobial Susceptibility Testing

2.2.1 Disc Diffusion

Antimicrobial susceptibility testing (AST) was performed using disc diffusion according to the British Society for Antimicrobial Chemotherapy (BSAC) methods (279). Three to five colonies of similar appearance were lightly touched with a 1µl loop. This was inoculated into a vial of 5ml sterile 0.9% saline (Oxoid Ltd, Basingstoke, UK) to achieve a 0.5 McFarland suspension. A sterile cotton-tipped applicator was used to spread the inoculum onto an Iso-Sensitest Agar

plate (Oxoid Ltd, Basingstoke, UK), first with a single streak down the centre of the plate, then zig-zagged across the plate 3 times, rotating the plate 60° in between. Antimicrobial Susceptibility Discs (Oxoid Ltd, Basingstoke, UK) were applied to the inoculated plate. Plates were then incubated 18 – 20 hours aerobically at 35°C. Zones of inhibited growth around the antibiotic disc were measured and used to determine resistance (R) , intermediate (I) or susceptible (S) according to the clinical breakpoints outlined by BSAC and as shown in Table 2.1 (280). For disc diffusion methods cefoxitin resistance is used as a surrogate marker for methicillin/oxacillin resistance (279,280).

METHOD		INTERPRETATION (mm)		
Antimicrobial	Disc Volume (µg)	S ≥	I	< R
Cefoxitin	10	27	22 – 26	21
Penicillin	1	*	*	*
Erythromycin	5	20	17 - 19	16
Gentamicin	10	20	-	19
Fusidic Acid	10	30	-	29
Clindamycin	2	26	23 - 25	22
Tetracycline	10	20	-	19
Ciprofloxacin	1	14	-	13
Linezolid	10	20	-	19
Rifampicin	2	30	24 - 29	23
Mupirocin	20	27	7 - 26	6

Table 2.1. Zone diameter breakpoints for disc diffusion method antimicrobial susceptibility testing of Staphylococci (280). S: susceptible, I: intermediate, R: resistant, *A heaped zone edge is used to indicate beta-lactamase mediated resistance.

2.2.2 Minimum Inhibitory Concentration

The minimum inhibitory concentration was tested using E-test © (BioMerieux, Marcy-l'Étoile, France) or M.I.C. Evaluator (Oxoid Ltd, Basingstoke, UK) strips which are impregnated with a gradient of known exponential concentrations of antibiotic. A 0.5 McFarland standard of the inoculum was prepared in sterile saline as above. This was used to inoculate an Iso-Sensitest Agar plate (Oxoid Ltd, Basingstoke, UK), first with a single streak down the centre of the plates, then zig-zagged across the plate 3 times, rotating the plate 60° in between. The E-test © or M.I.C.E strip was applied to the centre of the plate. Plates were then incubated 18 – 20 hours

aerobically at 35°C. For oxacillin testing, Columbia Agar Base with 2% Salt (CSA) plates (Oxoid Ltd, Basingstoke, UK) were used as the growth medium and these were incubated at 30°C for 24 hours before reading. Following incubation MICs were interpreted as the lowest concentration of antibiotic that inhibited the growth of test organism as shown in Table 2.2.

METHOD	INTERPRETATION (mg/L)		
Antimicrobial	S ≥	I	< R
Oxacillin	2	-	-
Penicillin	0.12	-	0.12
Erythromycin	1	2	2
Gentamicin	1	-	1
Vancomycin	4	-	4
Fusidic Acid	1	-	1
Clindamycin	0.25	0.5	0.5
Tetracycline	1	2	2
Ciprofloxacin	1	-	1
Teicoplanin	4	-	4
Linezolid	4	-	4
Rifampicin	0.06	0.12 – 0.5	0.5
Mupirocin	1	2 - 256	256

Table 2.2. MIC breakpoints for antimicrobial susceptibility testing of Staphylococci (280).

2.3 Storage

2.3.1 Bacterial Isolates

All isolates were frozen for long-term storage. Isolates were emulsified into a 200 µl solution of 20% glycerol (Sigma Aldrich, St. Louis, Missouri, US) and 80% sterile water (Oxoid, Basingstoke, UK). The emulsification was then pipetted into a Matrix™ 2D Barcoded Tube (Thermo Fisher Scientific, Massachusetts, US). They were then frozen at -80°C. Barcode, storage location and isolated identifier were archived in a biobank data.

2.3.2 DNA

For short-term storage DNA was stored in its elution buffer at -20°C.

2.4 Reference Strains

Two references strains were used in the research.

Staphylococcus epidermidis NCTC 11047 (ATCC 14990) from the National Collection of Type Cultures (Public Health England, London, UK) is the neotype strain of the species (*Staphylococcus epidermidis* (Winslow and Winslow 1908) Evans 1916) (4). For this strain, Whole-genome sequence data generated from PacBio sequences (Pacific Biosciences, California, US) and assembled and annotated by the Wellcome Sanger Institute are available at the European Bioinformatics Institute (EBI), European Nucleotide Archive (ENA): https://www.ebi.ac.uk/ena/data/view/ERS654925 consisting of 2,477,205 bp chromosome and 4 contigs (4866, 11,150, 19,915 and 24,597bp).

Staphylococcus epidermidis ATCC 12228 (Winslow and Winslow 1908) (281) is a highly characterised strain which is non-biofilm forming and commensal in origin (78,282). The whole-genome sequence is comprised of a 2,570,371 bp chromosome and five plasmids (282) and is available at the National Library of Medicine, National Center for Biotechnology Information (NCBI): https://www.ncbi.nlm.nih.gov/bioproject/PRJNA392321. The genome was sequenced at the Institute for Genome Studies, University of Maryland using PacBio (Pacific Biosciences, California, US), assembled using the Celera version 8.1 assembler and annotated with the NCBI Prokaryotic Genome Annotation Pipeline process (282).

2.5 Whole-Genome Sequencing

2.5.1 DNA extraction

DNA extraction was performed using the QuickGene DNA Tissue kit (DT-S) (Kurabo Industries Ltd., Osaka, Japan) and FastPrep-24 (MP Biomedicals, California, US) (Figure 2.1).

Fresh cultures from overnight incubation on CBA plates were used for chromosomal and mobile DNA extraction. A lysis solution of 480 µl LDT-01 lysis buffer (Kurabo Industries Ltd., Osaka, Japan), 20 µl EDT-01 Proteinase K solution (DT-S) (Kurabo Industries Ltd., Osaka, Japan) and 200 µl sterile saline (Oxoid Ltd, Basingstoke, UK) was added to Lysing Matrix-B tube (MP Biomedicals, California, US). A 5 µl loopful of bacterial isolate was emulsified in the solution. Mechanical disruption of bacterial cells was achieved processing inoculated tubes in a FastPrep-24 (MP Biomedicals, California, US) which homogenises the inoculum by rapid agitation (2 cycles of 6m/s for 40s). The homogenised solution was centrifuged at 16,100 rcf for 10 minutes to separate the DNA from the larger cellular components and silica beads. The

tubes were then heated at 70°C for 10 minutes to denature the DNA into single strands. Once denatured, the supernatant was mixed with 450 µl absolute ethanol to precipitate the DNA. 700 µl of precipitated DNA was pipetted into a CA cartridge (Kurabo Industries Ltd., Osaka, Japan). The extraction cartridge contains an ultra-thin polymer membrane which binds the DNA and allows other material to be removed separated and washed from other cellular components. Cartridges were processed in the QuickGene Mini80 (FujiFilm, Minato, Tokyo, Japan) where they underwent three washing steps with a solution of 350 µl WDT-01 Wash Buffer (Kurabo Industries Ltd., Osaka, Japan) and 350 µl absolute ethanol. QuickGene DNA Tissue kit (DT-S). Following washing the DNA was eluted with 200 µl CDT-01 elution buffer (Kurabo Industries Ltd., Osaka, Japan).

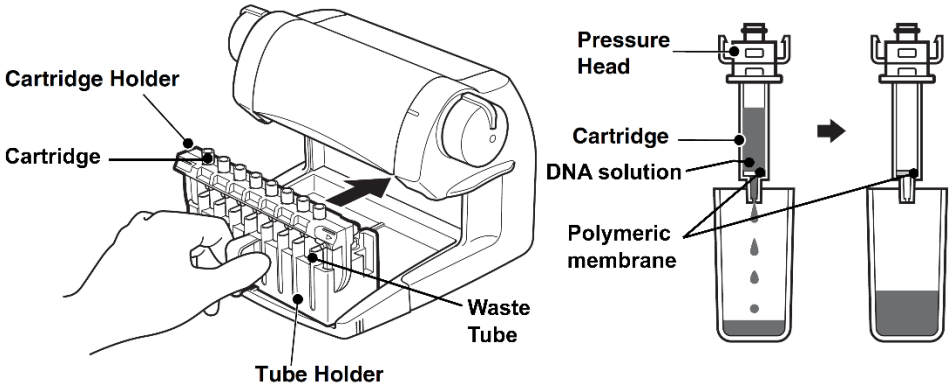


Figure 2.1. Use of the QuickGene Mini80 to extract bacterial DNA (adapted from the FujiFilm QuickGene Mini80 Nucleic Acid Isolation Device, User's Manual).

2.5.2 DNA Quantitation and Normalisation

Following extraction, the DNA was then quantitated and normalised to a concentration required by the sequencing library preparation protocol. Quantitation was performed using a fluorometric assay whereby a fluorometric dye which selectively binds dsDNA is mixed with the sample. Laser excitation (at 480nm) of the bound dye then generates a fluorescence which can be measured and compared against a standard curve of known DNA concentrations and their level of fluorescence.

For smaller batches, the Qubit 2.0 fluorometer (Invitrogen, California, US) was employed to measure DNA assayed with the Qubit dsDNA HS Assay Kit (Invitrogen, California, US). Large batches of DNA samples were prepared in a 96-well plate using Quant-iT PicoGreen dsDNA

Assay Kit (Invitrogen, California, US) and quantitated with an Infinite Pro 200 plate reader (Tecan, Männedorf, Switzerland).

In-house DNA preparation calculators were used to record DNA concentrations and calculate the volumes required to produce samples of a suitable concentration for DNA library preparation.

2.5.3 Illumina MiSeq DNA Library Preparation

Extracted DNA was prepared for Illumina sequencing by undergoing a process of library preparation using either the Nextera or Nextera XT DNA Library Prep kit (Illumina, California, US).

5 µl of 0.2ng/µl was added to 10 µl TD buffer. Then 5 µl of the transposome containing ATM mix (Illumina, California, US) was added and pipetted up and down to mix. This was heated at 55°C for 5 minutes then cooled to 10°C. Immediately after cooling, 5 µl of ND mix was added to stop the transposase reaction. This was incubated at room temperature for 5 minutes. To each sample 15 µl of the NPM polymerase mix was added alongside 10 µl of a combination of index primers. Table 2.3 shows the combinations of 5' and 3' index primers used for each MiSeq run.

Sample	Index Primer 1									Index Primer 2							
1	N701	T	A	A	G	G	C	G	A	S501	T	A	G	A	T	C	G	C
2	N702	C	G	T	A	C	T	A	G	S501	T	A	G	A	T	C	G	C
3	N703	A	G	G	C	A	G	A	A	S501	T	A	G	A	T	C	G	C
4	N704	T	C	C	T	G	A	G	C	S502	C	T	C	T	C	T	A	T
5	N705	G	G	A	C	T	C	C	T	S502	C	T	C	T	C	T	A	T
6	N706	T	A	G	G	C	A	T	G	S502	C	T	C	T	C	T	A	T
7	N701	T	A	A	G	G	C	G	A	S503	T	A	T	C	C	T	C	T
8	N702	C	G	T	A	C	T	A	G	S503	T	A	T	C	C	T	C	T
9	N703	A	G	G	C	A	G	A	A	S503	T	A	T	C	C	T	C	T
10	N704	T	C	C	T	G	A	G	C	S504	A	G	A	G	T	A	G	A
11	N705	G	G	A	C	T	C	C	T	S504	A	G	A	G	T	A	G	A
12	N706	T	A	G	G	C	A	T	G	S504	A	G	A	G	T	A	G	A

Table 2.3. Combinations of index primers and their nucleotide sequences used for MiSeq sequencing run of 12 samples. Nucleotides are coloured red (A and C) and green (G and T) as these are the colours of the fluorophore-labelled reversible terminators. Indexing must be set up so that the same position on all primers is not the same colour.

Characteristics of the PCR reaction and thermocycling program are shown in Table 2.4 and Table 2.5.

Reagent	Volume
Index 1	5 µl
Index 2	5 µl
Nextera PCR Master Mix	15 µl
Input DNA	25 µl
Total	50 µl

Table 2.4. PCR amplification reaction.

Temperature	Duration	Cycles
72°C	3 min	1
95°C	30 s	1
95°C	10 s	12
55°C	30 s	
72°C	30 s	
72°C	5 min	1
10°C	Hold	-

Table 2.5. PCR amplification thermocycling program.

A clean-up of the amplified reads was then performed to leave only fragments of DNA $\geq$ 100 bp. First, 25 µl of magnetic Agencourt AMPure XP beads (Beckman Coulter, California, US) was added to each sample and mixed well. This was incubated at room temperature for 5 minutes. The PCR plate was put onto a magnetic plate holder to hold the magnetic beads against the side of the plate well. All the supernatant was pipetted off. 200 µl of 80% ethanol was added to the well and incubated for 1 minute then pipetted off. This ethanol wash was repeated once more then left at room temperature for 15 minutes to air dry.

2.5.4 Illumina sequencing

Whole-genome sequencing was performed using an Illumina MiSeq at the Royal Sussex County Hospital, Brighton and by the Wellcome Trust Centre for Human Genetics (WTCHG), Oxford using an Illumina HiSeq4000. Each run of HiSeq sequencing included a control strain of *S. epidermidis* NCTC 11047 which was mapped against the assembled genome from the European Nucleotide Archive (ENA) as a quality control. Figure 2.2. provides an illustrative summary of Illumina sequencing.Illumina output files are in FASTQ format which is a text-based format that includes the nucleotide sequence of the read and its corresponding quality scores. The average read length across all genomes sequenced was 320 bp.

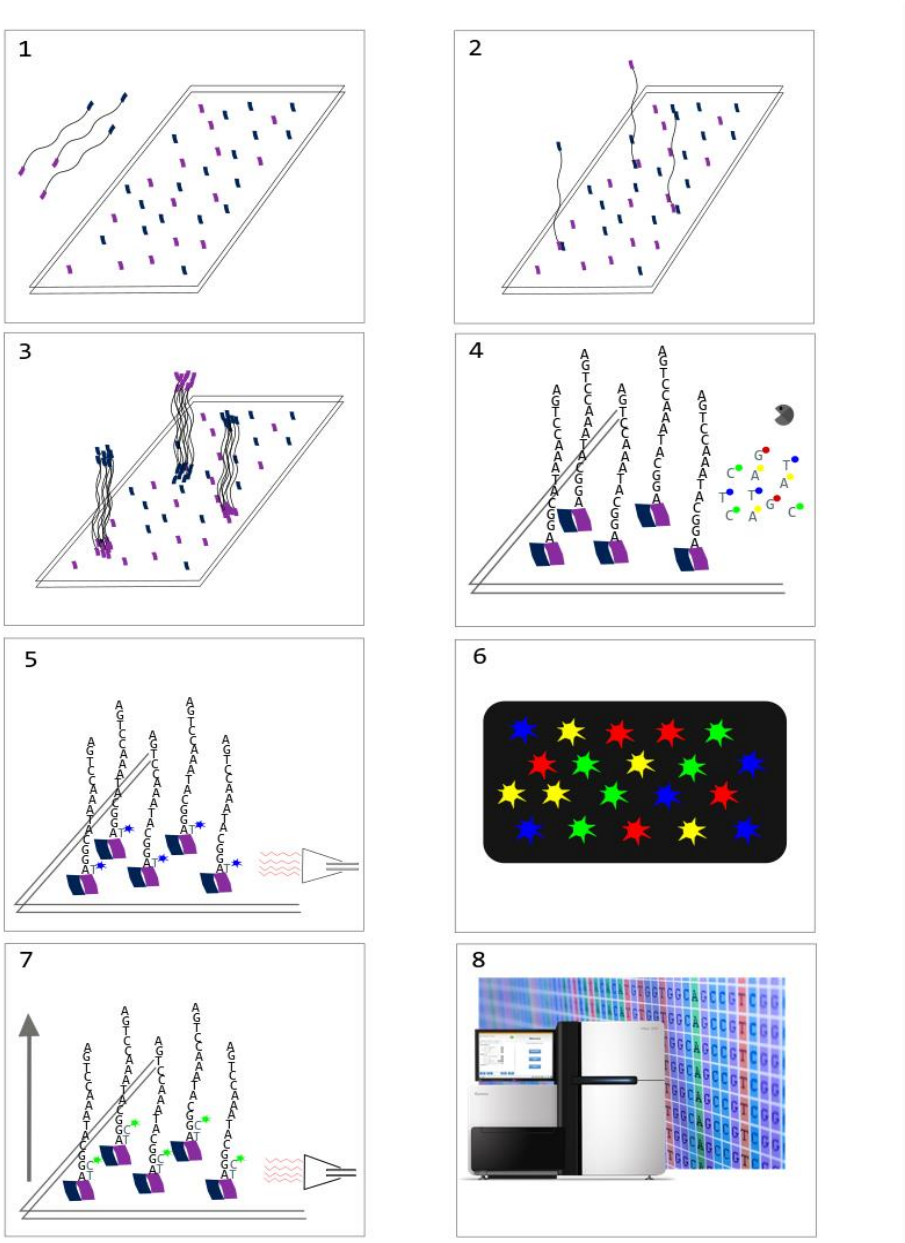


Figure 2.2. Cartoon summarising Illumina sequencing. 1) The genome library consisting of adaptor-bound short read DNA fragments is added to the flow-cell in the Illumina instrument. 2) The adaptor sequences on the DNA reads bind to those fixed on the flow-cell so that each DNA read is vertically positioned. 3) Identical clusters of each DNA sequence are generated. 4) The flow-cell is flooded with fluorophore-labelled reversible terminators, primers and DNA polymerase. 5) After a wash step only bound fluorophore-labelled reversible terminators remain which are complementary to the first nucleotide in the DNA read. 6) A camera focused down on the flow-cell captures an image of the laser-excited fluorophore, the wavelength and strength of fluorescence of which being specific to each nucleotide (A, T, C, or G). 7) Steps 4 – 6 are repeated until the full complementary strand of DNA has been synthesised and the sequence of nucleotides determined. 8) During post-run analysis, the Illumina instrument uses the synthesised, complementary DNA sequence to determine the original DNA read sequence. As the first and last 8-nucleotide sequences on the reads represent the Index Primers the instrument uses these to determine which read belongs to which sample.

2.6 Bioinformatic Processing

Raw sequence data (FASTQ files) were processed through a bespoke object-relational mapping-based (ORM) sequencing pipeline developed by the Modernising Medical Microbiology (MMM) Consortium, Nuffield Department of Clinical Medicine laboratories, University of Oxford (http://modmedmicro.nsms.ox.ac.uk/). The pipeline includes 2 processes by which bacterial genomes are assembled from short reads.

2.6.1 Mapping-based assembly

Reference-based mapping aligns the short-read fragments against a fully assembled reference genome. The MMM pipeline uses the Stampy algorithm that identifies candidate mapping locations in the reference genome, filters the candidates for best matches, then uses a fast gapped aligner to align the read to the reference genome according to quality scores and the presence of indels (283). Downstream analyses using an approximate Bayesian model apply quality scores to identify the presence of single-nucleotide and indel polymorphisms (283). For this study we used the genome of the *Staphylococcus epidermidis* ATCC 12228 as the reference genome for mapping-based assembly with the Stampy algorithm.

2.6.2 *De novo* Assembly

Additionally, whole genomes are also assembled from short reads *de novo* using Velvet v1.2.10 (284). The Velvet algorithm employs de Brujin graphs, which represent overlaps in sequences of symbols. Velvet firstly requires the short reads to be aligned as k-mers (sequentially overlapping sequences of nucleotides where *k* represents the length of the read and adjacent k-mers overlap by k - 1) (Figure 2.3) (284). Velvet implementation of de Brujin graphs then constructs contiguous sequences (contigs) from the last nucleotide from each overlapping k-mer (284).

Pipeline output files were in two different formats. FASTA format is a text-based format that contains the contiguous sequences generated during mapping-based or *de novo* assembly. Binary Alignment/Map (BAM) files are in machine readable binary format of reads aligned against reference sequences.

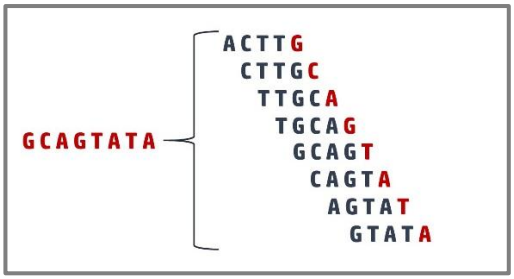


Figure 2.3. Velvet-based genome assembly. Overlapping short reads at a length of *k* (5 in this example) are aligned as k-mers at k-1 representing the original sequence (right). The correct nucleotide is determined as the final nucleotide in the k-mer (284).

2.6.3 Sequencing Data Quality

Sequencing of all isolates generated an average of 1843478 raw reads per genome with a G + C content of 32.5%. Velvet-based assembly generated an average of 82 contigs per genome. An average of 88.4% of the genomes mapped to the ATCC 12228 reference genome using Stampy with a coverage of 90.6%.

The MMM pipeline used in this study has an estimated false-positive rate of SNV (i.e. an error of a single nucleotide during assembly being mistaken for a point mutation) detection of 2.5 x 10^{-9} per nucleotide (285) or 0.00625 per genome of a 2.5 million bp genome. To ensure the quality of the whole-genome sequence data samples were sequenced in replicate at a rate of 5%. All replicated genomes had 0 SNVs difference. The reference strain NCTC 11047 was also included at a rate of 1% of samples. Each reference strain sequenced matched the complete reference genome of NCTC 11047 with 0 SNVs detected. Replicated and reference isolate genomes were removed from downstream analyses.

2.7 Analyses

2.7.1 Basic Local Alignment Search Tool

The Basic Local Alignment Search Tool (BLAST) searches databases of DNA or protein sequences (subject sequences) for identical or similar sequences submitted by the user (query sequences) (286,287). The BLASTn programme aligns nucleotide sequences (12). The tBLASTn programme aligns amino acid sequences to nucleotide sequences (12). Both programmes determine levels of identity which range from low similarity to perfects matches (12).

2.7.2 Multi-Locus Sequence Typing

Multi-locus sequence typing (MLST) compares the sequence variation of loci present in several house-keeping genes. The current MLST scheme was developed by Thomas *et al.* (2007) (246). As shown in Table 1.4, the scheme employs 7 loci and at the time of writing there were 782 sequence types (STs) of *S. epidermidis* comprising the profiles of 435 alleles (246,288).

Whole-genome sequence data was used to determine the ST of all isolates collected during this study. This was done *in silico* using a Bash script based on a script developed by Daniel Wilson (Figure 2.4) to interrogate the WGS data using BLASTn to identify alleles used for MLST. The allele profile was then used to determine the sequence type of the isolate.

```
---------------------------------------------------------------------
#!/bin/bash
# Usage:  sepimlst.sh genome_filename mlst_database_directory
# filters for nucleotide identity (perc_identity)=100, alignment length (length)=query length (qlen),
mismatches (mismatch)=0, gaps (gapopen)=0
#kevin.cole@bsuh.nhs.uk 11April2018
genomeList="$1"
dbdir="$2"
outputfile="$3"
#reads genomeList
export genomeNames=$(cat $genomeList)
for genome in $genomeNames
do
   zcat $genome > temp_genome.txt
   export arcCres=$(blastn -query $dbdir/Sepi/arcC.fas -subject temp_genome.txt -perc_identity 100
-outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export aroEres=$(blastn -query $dbdir/Sepi/aroE.fas -subject temp_genome.txt -perc_identity 100
-outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export gtrres=$(blastn -query $dbdir/Sepi/gtr.fas -subject temp_genome.txt -perc_identity 100 -
outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export mutSres=$(blastn -query $dbdir/Sepi/mutS.fas -subject temp_genome.txt -perc_identity
100 -outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk
'$4 == $3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export pyrRres=$(blastn -query $dbdir/Sepi/pyrR.fas -subject temp_genome.txt -perc_identity 100
-outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export tpiAres=$(blastn -query $dbdir/Sepi/tpiA.fas -subject temp_genome.txt -perc_identity 100
-outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export yqiLres=$(blastn -query $dbdir/Sepi/yqiL.fas -subject temp_genome.txt -perc_identity 100 -
outfmt '6 qseqid pident qlen length mismatch gapopen qstart qend sstart send evalue' | awk '$4 ==
$3 && $5 == 0 && $6 == 0' | awk '{print $1}' | sed 's/[^0-9]*//g')
   export MLSTres=$(awk '$2==ENVIRON["arcCres"] && $3==ENVIRON["aroEres"] &&
$4==ENVIRON["gtrres"] && $5==ENVIRON["mutSres"] && $6==ENVIRON["pyrRres"] &&
$7==ENVIRON["tpiAres"] && $8==ENVIRON["yqiLres"]' $dbdir/Sepi/ST.txt | awk '{print $1}')
   if [ -z $arcCres ]; then
            arcCres=0
   fi
   if [ -z $aroEres ]; then
            aroEres=0
   fi
   if [ -z $gtrres ]; then
            gtrres=0
   fi
   if [ -z $mutSres ]; then
            mutSres=0
```

```
    fi
    if [ -z $pyrRres ]; then
            pyrRres=0
    fi
    if [ -z $tpiAres ]; then
            tpiAres=0
    fi
    if [ -z $yqiLres ]; then
            yqiLres=0
    fi
    echo $genome $arcCres $aroEres $gtrres $mutSres $pyrRres $tpiAres $yqiLres $MLSTres >>
$outputfile
done
#to run: ./sepimlst.sh genomeList.txt ./ <outputFileName.txt>
--------------------------------------------------------------------------
```

Figure 2.4. Bash script developed to perform multi-locus sequence typing of *S. epidermidis* from whole-genome sequence data using BLASTn. Based on the script written by Daniel Wilson, Oxford.

2.7.3 Submission of Novel Sequence Types

Where previously undescribed STs were identified in this study these were checked manually using BLASTn (286). In each case data were then submitted to the Bacterial Isolate Genome Sequence database (BIGSdb) for availability at pubMLST.org (235,288). Isolate metadata, WGS FASTA files and allele FASTA files were uploaded to the database for approval and confirmation.

2.7.4 Constructing Haplotrees

To determine the genomic variation between colonies isolated from the same set of clinical specimens, SNVs and their position in the referenced-mapped genome were determined using an R-script developed by Professor Daniel Wilson (Nuffield Department of Medicine, University of Oxford). Output data were then used to generate haplotrees using a second R-script developed by Professor Daniel Wilson (Nuffield Department of Medicine, University of Oxford) whereby nodes of the tree represent genotype observation. The size of the node is proportional to the frequency of observation. Unobserved genotypes occurring between observed genotypes are also shown.

2.7.5 Detection of genetic elements associated with antimicrobial resistance

De novo assembled genomes of *S. epidermidis* (the subject sequences) were interrogated for the presence of the sequences in the two compiled FASTA files (the query sequences). This was done by using a Python script called Typewriter developed by Dr. Tanya Golubchik (https://github.com/tgolubch/typewriter) which runs BLASTn and tBLASTn (286,289–291).

Chapter 3

Measuring the diversity of *Staphylococcus epidermidis* in normal human carriage

3.1 Introduction

Staphylococcus epidermidis is a ubiquitous commensal member of the normal human flora that plays an important and complex role in the host microbial community, for example by inhibiting colonisation by more pathogenic species such as *S. aureus* (29,292). It can constitute up to 50% of colonies of commensal staphylococci isolated from human skin (293). Multiple strains of *S. epidermidis* may be carried on an individual's skin or mucosal surfaces at any one time (261,293). *S. epidermidis* may be present as a normal commensal at dry, moist and sebaceous areas of the skin as well in the throat and anterior nares (7,294,295). Previous research has described a relatively stable population structure which shows fluctuations over time (261). Some strains may become dominant and new strains may be introduced (132,261). Some acquisitions may be transient while others result in long-term colonisation (132). The ability of *S. epidermidis* to establish long-term human carriage is mediated by several factors (29,89,296). These include sodium ion/proton exchangers and osmoprotection transport systems which protect against high salt concentrations and osmotic pressure (29,89,296). In addition, protease SepA and ABC transporters mediate resistance to epithelial antimicrobial peptides by degradation and expulsion (296).

The multi-locus sequence typing scheme for *S. epidermidis* proposed by Thomas *et al.* (2007) uses the seven loci *arcC, aroE, gtr, mutS, pyrR, tpiA, yqiL* (Table 1.4) (246). To date, 782 individual sequence types from human, animal and environmental sources have been identified (297). Among these, ST2 is the most commonly isolated among clinical sample sets from across the globe (10,111,120). However, existing data are dominated by sample sets collected among hospitalised patients and cases of invasive infection. They under-represent the community, commensal *S. epidermidis* population (Table 1.3). *S. epidermidis* is ubiquitous both among healthy and hospitalised adults (298) however hospital stay, and prophylactic antibiotic treatment can result in the substitution of commensal community strains with more resistant, more virulent hospital-adapted strains (105,111,120,299).

Even among the detailed and comprehensive studies characterising the nosocomial strain diversity of *S. epidermidis* in humans, evidence has relied on low-discrimination techniques such as multi-locus sequence typing and pulsed-field gel electrophoresis (24,117,120,238,292,300–303). Consequently, intra-strain diversity of human carriage in this population has not been explored in detail. Measuring diversity at high resolution using single-nucleotide variants has been used to inform on carriage, transmission and infection of *S. aureus* (304). The effect of long-term asymptomatic carriage on the intra-strain diversity of *S.*

aureus has also been investigated using whole-genome sequencing (285) however, at present, no similar studies can be found involving *S. epidermidis*. Characterisation of *S. epidermidis* diversity within strains and over time during normal commensal carriage is needed to properly understand the population structure and dynamics of this organism. This knowledge would facilitate a population dynamics-based approach that could be used to discriminate between infecting and contaminating isolates of *S. epidermidis*.

3.2 Objectives

The work described in this chapter was designed to provide detailed characterisation of *S. epidermidis* carriage in healthy human volunteers regarding variation between body sites and over time by achieving three specific aims:

1) Characterise the multi-locus sequence type and single nucleotide variant diversity of *S. epidermidis* at multiple anatomical sites from multiple healthy donors.

2) Compare conventional methods used to determine *S. epidermidis* phylogeny with a sequence-based method derived from MLST data and determine the positions of carriage isolates from this study in phylogenetic trees derived using MLST-based and conventional methods.

3) Determine the mutation rate (molecular clock) for *S. epidermidis* during normal human carriage over a 12-month period.

3.3 Methods

3.3.1 Sampling

3.3.1.1 Ethical Approval

Ethical approval was granted for RGEC Ref No. 16/035/LLE, Diversity of Staphylococci Carried by Healthy People by the Brighton and Sussex Medical School Research Governance and Ethics Committee (RGEC) on the 17th November 2016.

3.3.1.2 Participants

The study set out to recruit 12 healthy individuals, 6 males and 6 females, aged between 18 and 60 and self-declared as healthy without a recent overnight hospital stay.

3.3.1.3 Sampling

Participants were given comprehensive Participant Information and Consent Sheet to read and sign and were assigned an anonymous identifier known only by themselves and me. Participant information (name, anonymous identifier, age, enrolment and swabbing dates) was recorded in a password protected database accessible only by me. Each participant consented to contribute seven self-taken superficial swabs (Table 3.1) every two months over a twelve-month period. The sampling sites were selected based on previous literature and were left and right axillae, anterior nares, outer cheeks and the throat (7,31,132,261,292–295,305,306). At each time point, participants were given a pack of 7 labelled swabs and a brief protocol for self-swabbing (Figure 3.1.)

Nasal

- Insert the swab inside the nostril so that just the cotton-tip is all the way in
- Rotate the swab about 5 times
- Return the swab back into its sterile container
- Repeat for the other nostril

Outer Cheek

- Move the swab back and forward over the **outer** cheek about 5 times in a zig-zag pattern
- Return the swab back into its sterile container
- Repeat for the other cheek

Throat

- Open your mouth wide
- Without touching your lips or cheek, insert swab the back of the mouth as best as possible
- Return the swab back into its sterile container

Armpits

- Please remove any top clothing to make this easier
- Move the swab back and forward over the armpit about 5 times in a zig-zag pattern
- Return the swab back into its sterile container

Figure 3.1. Mini protocol included in the swab pack.

3.3.2 Culture and Identification

3.3.2.1 Culturing from swabs

Swabs were inoculated in a 10ml 7% HPA Salt Broth (Oxoid Ltd, Basingstoke, UK) which enhances growth of halo-tolerant staphylococci whilst suppressing growth of other organisms (307). Broths were incubated overnight at 35°C. A 10µl loopful of the inoculated broth was spread onto a Columbia CAP Staph/Strep Selective Agar & Horse Blood (Oxoid Ltd, Basingstoke, UK) which is selective for Gram-positive organisms. The CAP agar plates were incubated aerobically at 35°C for up to 48 hours.

3.3.2.2 Isolation and Identification of *S. epidermidis*

Colonies of *S. epidermidis* were initially selected according to their visual appearance on Columbia Cap Selective Agar (Oxoid Ltd, Basingstoke, UK) as described in Section 2.1.2 Identification of *S. epidermidis.* Individual candidate colonies were then sub-cultured on to a Columbia Blood Agar (CBA) plate (Oxoid Ltd, Basingstoke, UK). Following incubation, colonies were identified to species using MALDI-ToF MS (Bruker Daltonik GmbH, Leipzig, Germany), (see Chapter 2: Materials and Methods). Based on experience within my research group and the published literature of SNP-based assessment of *S. aureus* diversity (285,308), and balancing sensitivity to detect low frequency variants against the cost of repeatedly sequencing each strain, a maximum of eight individual colonies of *S. epidermidis* from each cultured swab were selected for further analyses.

3.3.3 Antimicrobial Susceptibility Testing

3.3.3.1 Disc Diffusion

A crude method for distinguishing strains of *S. epidermidis* is to compare antimicrobial susceptibility patterns (the strain's "antibiogram"). The panel of 13 antimicrobials used was the same employed in standardised antimicrobial susceptibility testing (AST) of clinical isolates of staphylococci at the Royal Sussex County Hospital, Brighton. AST was performed using the disc diffusion methods according to the British Society for Antimicrobial Chemotherapy (BSAC) (279) which are described in Chapter 2: Materials and Methods. All selected colonies of *S. epidermidis* underwent AST to determine their antibiogram.

3.3.4 Selection of Isolates for Whole-Genome Sequencing

Random selection of isolates for sequencing was achieved by listing all of the isolates in an Microsoft Excel spreadsheet. A column was added to the list and the RAND() function was applied which generates random numbers between 0 and 1. Those with an odd number were selected out for WGS. This was repeated until sufficient isolates had been selected.

3.3.5 Whole-Genome Sequencing

Whole-genome sequencing was performed on an Illumina HiSeq4000 at the Wellcome Trust Centre for Human Genetics using methods described in Chapter 2: Materials and Methods.

3.3.6 Analyses

3.3.6.1 Multi-locus sequence typing

Multi-locus sequence typing was performed *in silico* from WGS data using a BLASTn-based method (see Chapter 2: Materials and Methods).

3.3.6.2 Minimum-spanning tree

To provide a visual representation of MLST-based phylogeny for *S. epidermidis* a minimum-spanning tree (MST) was constructed using the PHYLOViZ software (309). PHYLOViZ employs goeBURST, an implementation of the eBURST algorithm which constructs a phylogenetic tree based on allele profiles (234,310). The eBURST algorithm can cluster multi-locus sequence types according to allele variants (234,310). Single-locus variants are separated by a one differing allele, double-locus variants are separated by two differing alleles and triple-locus variants are separated by three differing alleles (234,310). The global optimal implementation of eBURST, goeBURST allows a global heuristic selection of the founding ST when clonal complexes are defined (310). Minimum-spanning trees were constructed of the 782 MLST allele profiles available at the time of writing and of the sequence types detected from isolates obtained during this study (see Chapter 2: Materials and Methods).

3.3.6.3 Maximum likelihood trees

To investigate the evolutionary diversification of *S. epidermidis* isolates maximum likelihood (ML) trees were constructed. The maximum likelihood tree method of phylogenetic analysis

represents the genetic sequences of individual organisms as the leaves of a tree. Single-nucleotide variants in the sequence are analysed so that the most likely points of evolutionary diversification between organisms can be represented by the points at which tree branches diverge . The branch lengths correspond to the amount of evolutionary separation (234,310). ClonalFrameML is a computational tool which detects and accounts for recombination events during ML tree construction (311). ML trees were created using a Python script generously shared by Dr. David Eyre (Nuffield Department of Medicine, University of Oxford) which uses PhyML (312) to construct phylogenetic trees and incorporates ClonalFrameML into the analysis.

3.3.6.4 Fixation index

To estimate the influence of anatomical site on genomic diversity of colonising populations of *S. epidermidis* a fixation index estimator was employed. Hudson's Fixation Index estimator can be used to measure the difference between diversity within a subpopulation against the diversity between subpopulations to determine the impact of recombination on population structure (313). I employed the formula,

$$F_{ST} = \frac{Hw - Hb}{Hb}$$

where *Hw* represents the mean number of pairwise distances within a subpopulation and *Hb* represents the mean number of pairwise distances between subpopulations (313). For this study, subpopulations were defined by location, either within an anatomical site on a participant or within a participant. A value closer to one suggests a more clonal population structure with selective sweeps causing little diversity within a subpopulation but high diversity between subpopulations (314). A value closer to 0 suggests recombination between subpopulations from different locations may help drive diversity.

3.4 Results

3.4.1 Sampling

Eight healthy participants were recruited, and swabs were collected every two months. However, due to the high diversity of strains detected by antibiogram at the first time point it was decided to focus on isolates collected from the first five participants at time points 0 and 12 to perform an in-depth investigation of within-host diversity. The five participating

volunteers comprised two females and three males with a median age of 36 (Table 3.1). No chronic health condition or recent protracted hospital stay were declared by any participant. One participant withdrew from follow up at month 6. Therefore, isolates of *S. epidermidis* from swabs taken at month 0 from five of the participants and from swabs taken at month 12 for four of the participants were analysed. Hereafter months 0 and 12 will be referred to as time point 0 and time point 12.

Cultured isolates from a total of 63 swabs (seven anatomical sites during nine swabbing events) were ultimately used for the analysis presented here. From these a total of 306 colonies of *S. epidermidis*, 163 from time point 0 and 143 from time point 12 were selected (Table 3.2). All 306 colonies were confirmed as *S. epidermidis* using MALDI-ToF (see Chapter 2: Materials and Methods) and underwent antimicrobial susceptibility testing according to the BSAC methods (data not shown). A maximum of eight colonies of *S. epidermidis* were selected from each cultured swab though some cultured swabs grew less than eight. The throat was the anatomical site that grew the least number of *S. epidermidis* colonies. Only 3/9 throat swabs cultured any *S. epidermidis* and from these only two grew at least eight colonies. In total 18 colonies were isolated from the throat. The right axilla and right nostril grew the next lowest number of colonies (n = 41). Whilst all other sites grew *S. epidermidis* in at least one participant from at least one time point not all sites consistently yielded *S. epidermidis* at both time points. No participant grew eight or more colonies from every site.

Participant	Gender	Age	Time Point (month) 0	Time Point (month) 12
F5	Female	35	02/12/2016	8/12/2017
F7	Female	29	18/11/2016	N/A
M4	Male	48	18/11/2016	16/11/2017
M7	Male	36	06/01/2017	23/01/2018
M8	Male	57	25/11/2016	10/01/2018

Table 3.1. Study participants and sampling time frame

ANATOMICAL SITE	PARTICIPANT	TIME POINT (Month) 0	12	TOTAL	PARTICIPANT	TIME POINT (Month) 0	12	TOTAL
L Axilla		8	5	13		0	8	8
R Axilla		8	0	8		0	8	8
L Cheek		8	8	16		0	8	8
R Cheek	F5	8	0	8	M7	2	8	10
L Nasal		8	0	8		8	8	16
R Nasal		8	8	16		0	4	4
Throat		0	0	0		8	8	16
Total		**48**	**21**	**69**		**18**	**52**	**70**
L Axilla		8	NA	8		7	8	15
R Axilla		8	NA	8		7	8	15
L Cheek		7	NA	7		8	8	16
R Cheek	F7	5	NA	5	M8	8	8	16
L Nasal		8	NA	8		0	7	7
R Nasal		8	NA	8		0	0	0
Throat		0	NA	0		0	0	0
Total		**44**	**NA**	**44**		**30**	**39**	**69**
L Axilla		2	8	10		25	29	54
R Axilla		1	1	2		24	17	41
L Cheek		4	6	10		27	30	57
R Cheek	M4	1	8	9	**All**	24	24	48
L Nasal		8	0	8	**Participants**	32	15	47
R Nasal		5	8	13		21	20	41
Throat		2	0	2		10	8	18
Total		**23**	**31**	**54**		**163**	**143**	**306**

Table 3.2. The number of colonies of *S. epidermidis* that were isolated from each participant, time point and anatomical site.

3.4.2 Whole-Genome Sequencing Selection

To moderate WGS processing costs, 192 (192/306) colonies (92 from time point 0 and 100 from time point 12) were randomly selected for whole-genome sequencing analysis (Table 3.3). Additionally 9/192 (5%) colonies were sequenced in duplicate, in accordance with MMM and Wellcome Trust Centre for Human Genetics protocols, to detect false-positive variant calling which is estimated to occur at a rate of approximately 2.5×10^{-9} per nucleotide using the MMM pipeline (285). The nine duplicates differed by 0 SNVs from the originals and were excluded from downstream analysis (Table 3.4). Table 3.5 summarises the quality data for the 192 isolate sequences.

	F50AxRm1.2	F50ChRm4.2	M40NaRm3.2	M412ChRm2.2	M412NaRm7.2	M712AxLm5.2	M712Thm6.2	M80AxLm4.2	M812ChRm5.2
F50AxRm1	0	55481	54736	54757	54764	9241	54744	54667	56758
F50ChRm4	55459	0	7776	7782	7790	54533	8346	8458	28297
M40NaRm3	54727	7775	0	191	195	53977	7175	8393	28520
M412ChRm2	54741	7784	192	0	18	53995	7192	8417	28532
M412NaRm7	54759	7794	196	18	0	54003	7205	8426	28540
M712AxLm5	9267	54559	53977	53994	53999	0	53998	54094	55410
M712Thm6	54737	8339	7175	7193	7195	53995	0	9052	28261
M80AxLm4	54642	8449	8394	8405	8413	54074	9040	0	28959
M812ChRm5	56757	28301	28509	28505	28511	55418	28243	28938	0

Table 3.3. Pairwise distance matrix showing the single-nucleotide variation between original and duplicate samples that were whole-genome sequenced in duplicate to ensure accuracy of data. 9/192 samples were sequenced twice and a pairwise comparison of SNV differences in the genome were measured. There were 0 SNVs between original (rows) and duplicate (columns) samples when mapped to reference genome ATCC 12228 therefore there was limited chance of SNVs being falsely called.

ANATOMICAL SITE	PARTICIPANT	TIME POINT 0	12	TOTAL	PARTICIPANT	TIME POINT 0	12	TOTAL
Left axilla		4	2	6		0	6	6
Right axilla		4	0	4		0	3	3
Left cheek		4	7	11		0	5	5
Right cheek	F5	4	0	4	M7	2	4	6
Left nasal		4	0	4		4	6	10
Right nasal		4	7	11		0	2	2
Throat		0	0	0		4	5	9
Total		**24**	**16**	**40**		**10**	**31**	**41**
Left axilla		4	NA	4		4	5	9
Right axilla		5	NA	5		4	7	11
Left cheek		4	NA	4		5	5	10
Right cheek	F7	4	NA	4	M8	4	7	11
Left nasal		5	NA	5		0	5	5
Right nasal		4	NA	4		0	0	0
Throat		0	NA	0		0	0	0
Total		**26**	**NA**	**26**		**17**	**29**	**46**
Left axilla		2	8	10		14	21	35
Right axilla		1	0	1		14	10	24
Left cheek		2	4	6		15	21	36
Right cheek	M4	0	6	6	All	14	17	31
Left nasal		4	0	4		17	11	28
Right nasal		4	6	10		12	15	27
Throat		2	0	2		6	5	11
Total		**15**	**24**	**39**		**92**	**100**	**192**

Table 3.4. The number of colonies of *S. epidermidis* that underwent whole-genome sequencing from each participant, time point and anatomical site.

Paramater	Range	Median (IQR)
Reads (n)	1.96 - 4.98 x 10^6	3.54 x 10^6 (3.2 – 4.03 x 10^6)
Mapped (n)	1.61 - 4.27 x 10^6	3.15 x 10^6 (2.79 – 3.54 x 10^6)
Mapped (%)	73.2 - 95.7	87.5 (85.3 – 91.4)

Table 3.5. Quality control data for 192 carriage isolate genome sequences

3.4.3 *In silico* Multi-Locus Sequence Typing

By extracting from the WGS data sequences for the seven house-keeping genes used for *S. epidermidis* MLST it was possible to assign all the sequenced isolates to MLSTs *in silico* (246,288). A total of 32 different STs were identified from the five participants (Table 3.6). Four STs were detected from more than one participant: ST59 (4/5), ST73 (3/5), ST190 (2/5) and ST297 (2/5). ST73 was the most frequently detected; it accounted for 20/192 (10.4%) of the colonies and was isolated from 4/5 participants. The next most frequently isolated was ST153 (18/192 (9.4%)), although all colonies were from the same person; M4.

The number of different STs identified from each participant at each time point ranged from 2 - 10 (median = 7). From one participant (M4) only five STs were identified across both time points with two STs (ST153 and ST384) accounting for 29/39 (74%) isolates. Isolates from participants F5, M7 and M8 contained more diverse STs Notably, M4 was the participant with the least number of different sequence types. In participant M7 only two different STs were detected at time point 0 but 10 different STs were detected at time point 12.

Over one third of the STs identified in these five participants (12/32) were novel MLST profiles, of which half (6/12) were the result of novel alleles. The twelve profiles and six allele sequences were submitted and archived in the pubMLST BIGSdb database (Table 3.7) (297). All twelve of the novel STs were isolated from individual participants and each participant yielded at least one novel ST.

PARTICIPANT	ST	TIME POINT 0	12	TOTAL	PARTICIPANT	ST	TIME POINT 0	12	TOTAL
	32	8	0	8		17	0	1	1
	59	0	1	1		59	3	2	5
	73	6	2	8		73	7	1	8
	190	3	0	3		130	0	6	6
F5	723	1	5	6		190	0	1	1
	766	3	0	3	M7	457	0	7	7
	767	1	0	1		558	0	3	3
	774	2	0	2		**765**	0	2	2
	778	0	8	8		**768**	0	1	1
	d	7	4	9		**773**	0	7	7
	73	3	NA	3		d	2	10	10
	297	8	NA	8		19	0	4	4
	487	4	NA	4		57	9	0	9
F7	543	1	NA	1		59	0	3	3
	769	7	NA	7		73	0	1	1
	775	2	NA	2	M8	87	5	5	10
	776	1	NA	1		297	0	5	5
	d	7	NA	7		329	0	8	8
	7	0	3	3		**770**	3	3	6
	89	5	0	5		d	3	7	8
M4	153	6	12	18					
	384	3	8	11					
	777	1	1	2					
	d	4	4	5					

Table 3.6. The number of colonies of each *S. epidermidis* ST isolated and analysed from each participant and time point. **d** denotes the number of different STs detected in a single participant. Novel STs are highlighted bold.

Novel MLST Types	Allelic Profile							Specimen
	arcC	*aroE*	*gtr*	*mutS*	*pyrR*	*tpiA*	*yqiL*	
765	12	48	5	5	10	16	12	Carriage
766	12	48	5	5	10	16	21	Carriage
767	28	48	5	5	10	40	31	Carriage
768	12	48	9	9	40	16	21	Carriage
769	1	1	2	1	2	1	45	Carriage
770	1	13	7	6	2	5	1	Carriage
773	1	1	7	1	**60**	5	14	Carriage
774	1	1	1	2	2	1	**63**	Carriage
775	2	1	6	2	2	1	**64**	Carriage
776	**69**	19	4	4	9	10	2	Carriage
777	8	7	12	4	12	**53**	2	Carriage
778	2	1	1	1	2	**54**	1	Carriage

Table 3.7. Novel MLST allele profiles submitted and accepted into the pubMLST BIGSdb *Staphylococcus epidermidis* database during this study. Novel alleles are highlighted in bold.

3.4.4 Minimum-Spanning Trees of Study Isolate Sequence Types

To better visualise the diversity of STs identified in the healthy participants minimum spanning trees (MST) of the sequence types were constructed (Figure 3.2). The trees demonstrate that individual participants carried STs that are phylogenetically diverse. Although ST57 was only isolated from one participant (M8) and at one time point (0) it forms the central node of the largest tree as the proposed founder. ST57 has nine branches radiating from it including three single locus variants and four double locus variants. ST19 differed from ST57 by 7/7 loci differences and therefore comprised a separate tree. The twelve novel sequence types were distributed across the trees suggesting that they are related to several previously known sequence types. However, 2/4 novel STs from participant F5 and 2/3 novel STs from participant M7 were clustered together in a group on the longest branch of the largest tree. Sequence type 457 and ST558 are situated on the distant edge of this cluster suggesting that the cluster of novel STs is ancestral to ST457 and ST558.

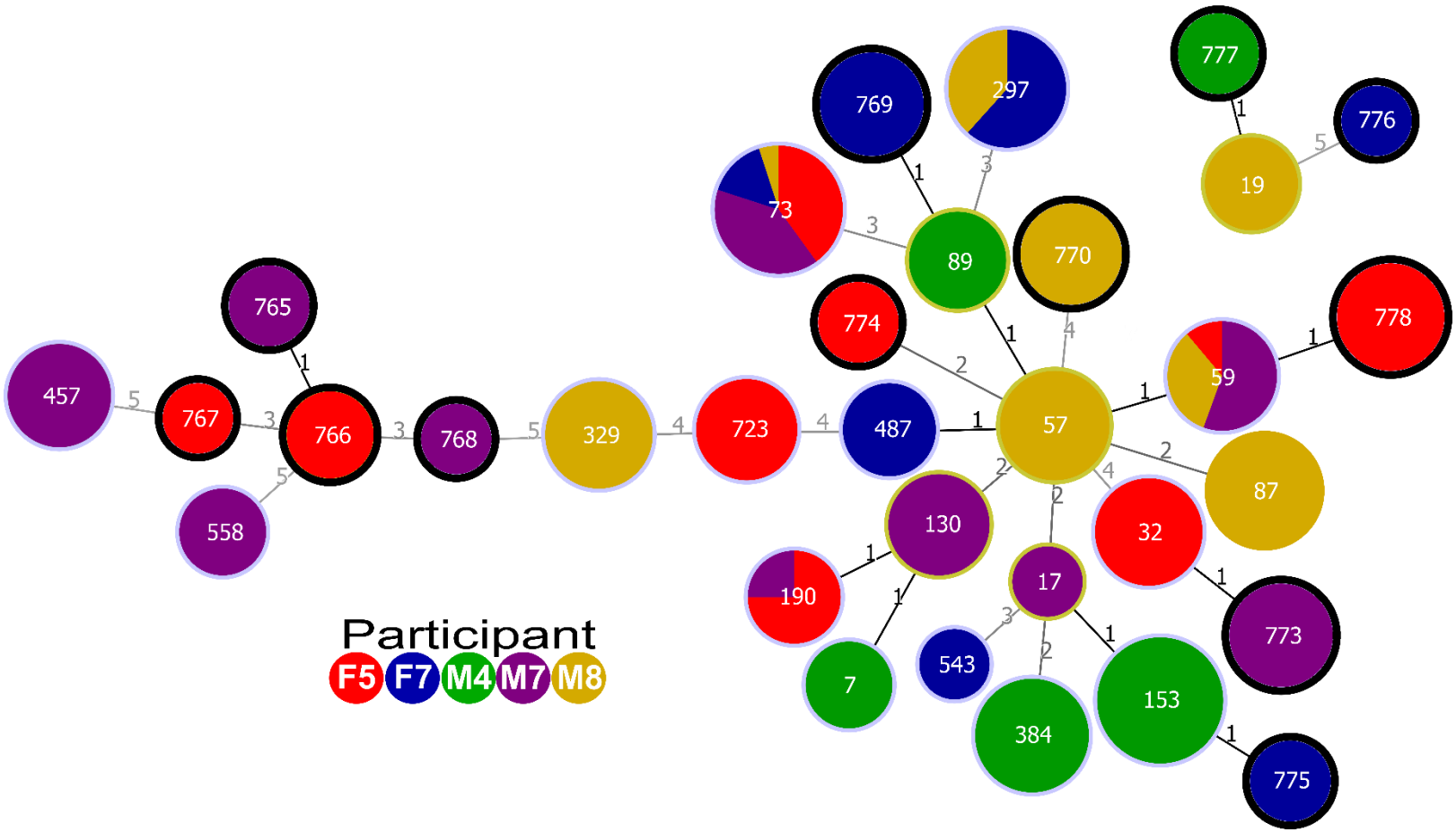


Figure 3.2. Minimum spanning trees of all carriage isolates according to multi-locus sequence type allele profile. Joining lines are labelled with the number of locus variants between STs. Nodes are labelled according to ST and coloured according to participant from which the ST was isolated. Node sizes are proportional to the frequency of ST isolation. Nodes circled in black represent novel STs.

3.4.5 Evaluation of healthy carrier diversity of *S. epidermidis* in context of the wider *S. epidermidis* population

At the time of writing, the MLST scheme for *S. epidermidis* comprises 782 STs comprised of 435 alleles (246,288). To place the isolates of *S. epidermidis* identified in this study in the wider context of the *S. epidermidis* STs using different phylogenetic techniques profile-based, minimum spanning method and sequence-based, maximum-likelihood method trees were constructed. To construct the minimum-spanning tree the current database of *S. epidermidis* STs was download from pubMLST.org (315). The database was uploaded to the PhyloViz software and the goeBURST algorithm was employed to construct the tree based on the STs being joined by the most closely related allelic profiles (Figure 3.3). Clonal complexes were formed of sequence types that were single-locus variants (6/7 loci were identical). To make the ML tree the sequences of the seven loci alleles comprising each ST were concatenated to give a single sequence per ST. The ML tree was constructed in ClonalFrameML. Sequence type 5 was used as the mapping reference as this is the proposed ancestor of the largest known clonal complex (117). Interactive Tree of Life (iTOL) v4.2.4 (316) was used to root the tree to ST5 (the blue node circled in black) and colour the tree nodes according to the clonal complexes assigned by goeBURST above.

The MST demonstrated that the four largest clonal complexes (CCs) were comprised of 466, 40, 16 and 10 STs respectively (Figures 3.3 and Table 3.8). Thirty-one other clonal complexes were comprised of smaller numbers of STs including 19 pairs. There were 164 singletons. In-keeping with the current understanding of how diversity within *S. epidermidis* has evolved, the MST places ST5 at the centre, as the assumed ancestor, of the largest clonal complex (117). Small clonal complexes and singleton sequence types are distributed across MST suggesting intermediate sequence types remain undetected. Eighteen (56%) of the STs isolated from the participants belonged to CC5, 1 ST from CC365, 2 STs from CC227 and 1 ST from CC193. The remaining 10 STs belonged to other, smaller CCs or were singletons. Sequence types isolated from the different participants were distributed across the tree suggesting that the diversity of *S. epidermidis* observed spans that seen in the human population in general. This in turn suggests that *S. epidermidis* evolution and diversification is a constant process and that has been occurring within the colonised human population for many years. However, the presence of the same ST on multiple participants also accords with the current literature that certain strains are more commonly carried (10,238). Recent work suggesting global spread of multidrug resistant strains demonstrates that dissemination of successful strains may also occur on a relatively short timescale (182).

In keeping with the MST, the ML tree showed that clonal complexes, especially CC227 but also most CC365 and CC193 isolates, generally sit together on branches of the ML tree although isolates of CC5 are distributed right across the tree including the most distant branches (Figure 3.4). This demonstrates that both methods generally concur at showing the evolutionary divergence are distributed across the tree suggesting intermediate STs remain undetected. A striking feature of the maximum likelihood tree is the isolated position of the single sequence type ST781 at the end of a long branch. However, this isolate lacks an entire gene (*arcC*); one of the housekeeping genes on which the MLST scheme is based. The isolated position of ST781 on the ML tree is an artefact caused by the loss of a single gene composed of 472 nucleotides. A minimum spanning tree more accurately represents the relatedness of ST781 to other sequence types and shows ST781 to be only one profile distant from ST2 because apart from the loss of *arcC*, all loci are shared by both sequence types.

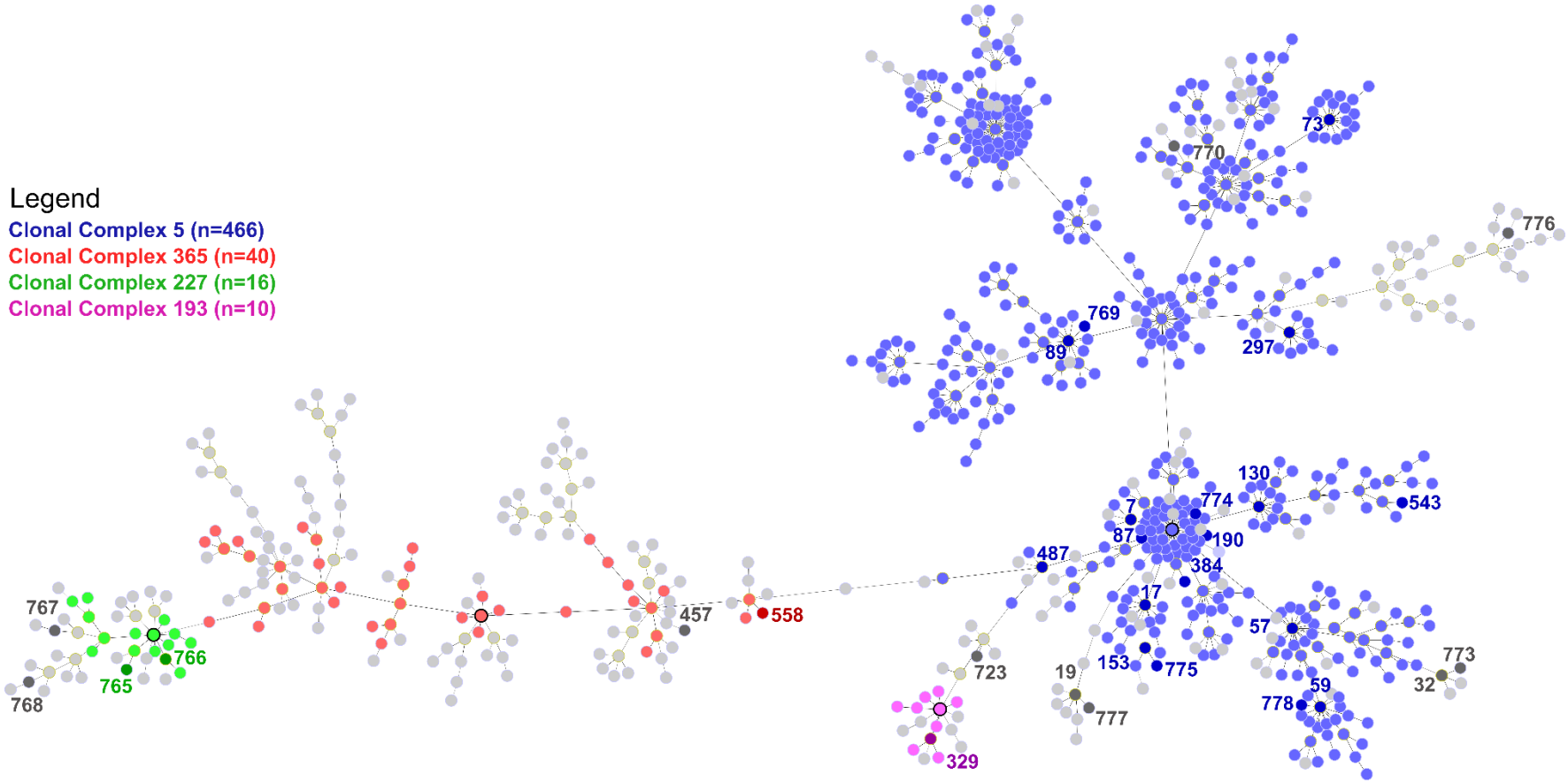


Figure 3.3. goeBURST minimum-spanning tree of 782 *S. epidermidis* multi-locus sequence type profiles. Clonal Complexes were grouped by single-locus variants. The four largest clonal complexes are coloured and named according to the proposed founding STs which are circled in black. Smaller clonal complexes and singletons are grey. Sequence types isolated from the carriage study participants are shaded darker and are labelled according to sequence type.

Number of STs in Clonal Complex	Frequency
466	1
40	1
16	1
10	1
6	2
5	1
4	4
3	5
2	19

Table 3.8. goeBURST algorithm computation of clonal complexes grouped by single locus variant from 782 MLSTs. 164 singletons were not included.

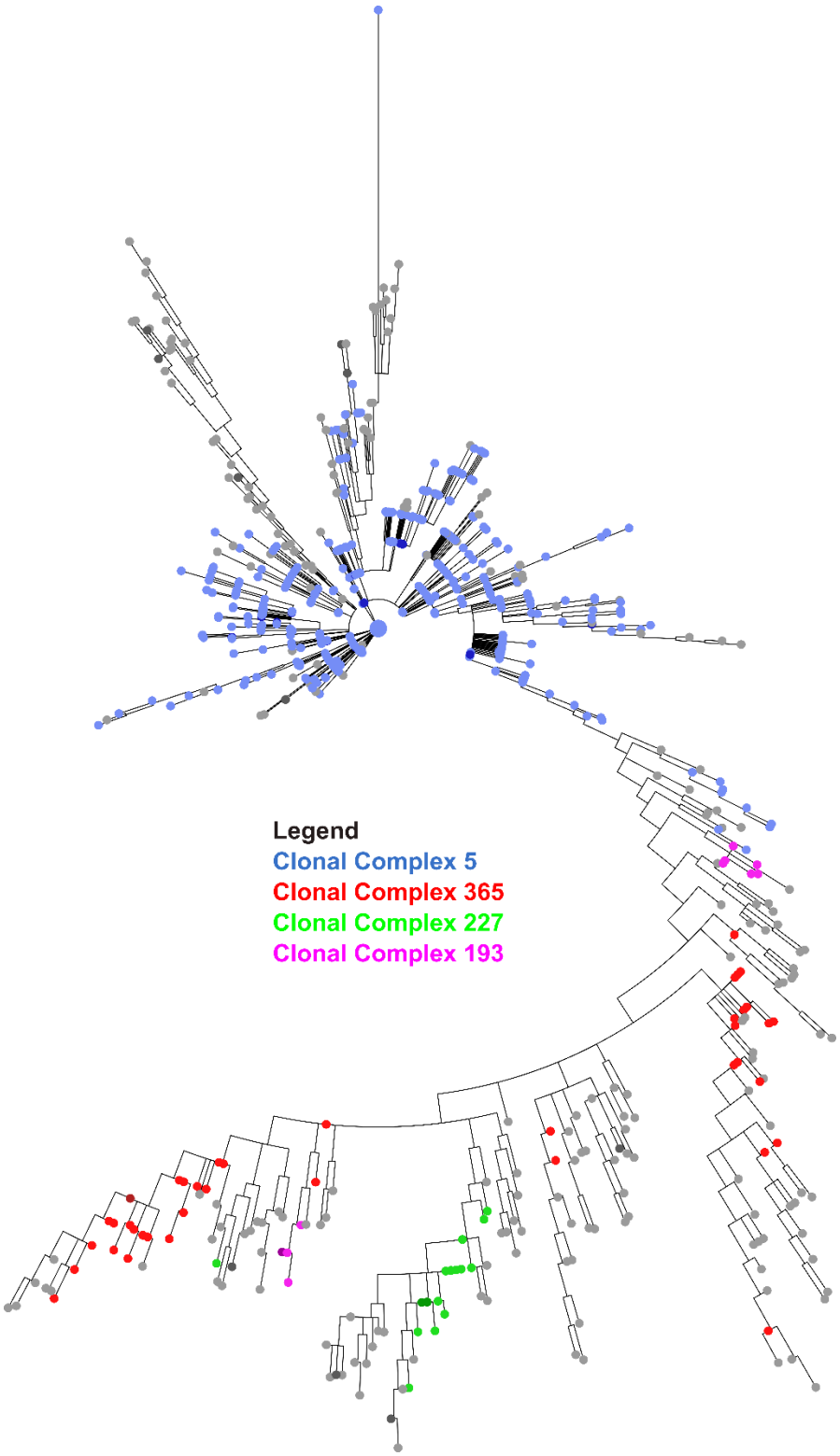


Figure 3.4. Maximum likelihood tree of concatenated *S. epidermidis* MLST alleles according to sequence type. Concatenated alleles of 782 STs were mapped to those of ST5 (the central large node) and the tree was constructed using ClonalFrameML. Branch lengths were cube-rooted to reduce relative length. The tree was then rooted to ST5. The four largest clonal complexes are coloured, smaller clonal complexes and singletons are grey. Sequence types isolated from the carriage study participants are shaded darker.

3.4.6 Whole-Genome Sequencing Pairwise Distance Analysis

3.4.6.1 Comparison between *S. epidermidis* carriage with *S. aureus* carriage

As previously discussed, in the field of *S. aureus* research, assessment of genome-wide single-nucleotide differences has allowed discrimination of closely related isolates in studies of transmission and pathogenesis (317,318). A pragmatically selected threshold of 40 SNVs to assign closely related 'subtypes' (equating to around 10 years of evolutionary time for *S. aureus*) has proven robust in transmission studies. *S. aureus* carriage, however, can be detected in only a proportion of the human population at any time and often only a single type or subtype can be detected in an individual. In contrast, *S. epidermidis* carriage appears to be universal and people are often colonised by multiple lineages simultaneously. To investigate the relatedness of the carriage isolates with higher resolution than MLST, pairwise distances of whole-genome sequence data were analysed. Measuring SNV differences between isolates is a more accurate method of determining relatedness than MLST as isolates of the same sequence type may in fact be distantly related genomically. This is particularly true for successful STs that have spread and diverged over time. The maximum within-participant diversity of 40 SNVs between 12 colonies detected by Golubchik and colleagues (2013) was applied by Price *et al*. (2016) as a threshold to assign *S. aureus* carriage isolates into subtypes to define acquisition (285,317). Furthermore, for this study the number of colonies isolated per sequence type was eight or less meaning a measurement of the maximum pairwise distance and mean pairwise distance would be a better indicator of relatedness. Therefore, to investigate the potential of assigning *S. epidermidis* isolates to subtypes several thresholds were investigated. To detect and remove genomic differences resulting from recombination events data output from ClonalFrameML was used (see Chapter 2: Methods). Only colonies isolated at the same time point were analysed by pairwise distance to remove the potential effect of temporal mutation rates of *S. epidermidis* on the analysis of genomic diversity.

Four comparisons were made to understand the impact of diversity within participants and within sequence types.

1) pairwise distances between colonies of the same and of different sequence types,
2) pairwise distances between colonies isolated from the same and from different participants,
3) pairwise distances between colonies isolated from the same and different anatomical sites on the same participant,
4) pairwise distances of all colonies regardless of participant or sequence type.

First the median values of the above comparisons were measured. Table 3.9 shows that colonies of the same ST had median diversities in the tens and hundreds whilst colonies of different sequence types had median diversities in the thousands. Overall median diversity was more affected by ST ($p=0.005$) than participant ($p=0.49$) by two-sample t-test. However, restricting ST diversity to within-participant only makes a log fold difference to the median SNVs. If the participant diversity is restricted to within-ST only there is a 2 log-fold difference in median SNVs. These data illustrate that lineages were more conserved across the participants than across STs.

ST	Participant	Median (IQR) (SNVs)	Range (SNVs)
Same	Same	14 (2 – 21)	0 - 550
Same	Different	479 (112 – 500)	34 - 1095
Different	Same	4517 (2399 – 11411)	60 – 13,661
Different	Different	3859 (2331 – 8728)	151 – 15,384
All	All	3647 (2253 – 8701)	0 – 15,384

Table 3.9. Median, interquartile range and range of pairwise distances between colonies isolated from the same time point.

3.4.6.2 Limits of isolate relatedness

Measuring SNV differences between all colonies in the sample set yielded 36,672 pairwise differences of which 18,272 were from the same time point. Pairwise distances from the same time point were plotted to give a snapshot of diversity as shown in Figure 3.5. Pairwise distances from different time points were removed from the data as mutations occurring over time may artificially increase diversity.

3.4.6.3 Within-sequence type diversity

Across the whole data set diversity within sequence type was markedly lower than between sequence types (Figure 3.5) with median of 17 SNVs (IQR = 3 – 36 SNVs), and 2354 SNVs (IQR = 3990 – 8793 SNVs) respectively . Within-ST diversity was lower within participants than between participants.

The minimum pairwise distance between colonies of the same sequence type isolated from different participants was 34 SNVs (Figure 3.5). This, alongside the high maximum number of 550 SNVs within participants suggests that there may have been a recent transmission or common ancestor between participants.

3.4.6.4 Between-sequence type diversity

The minimum pairwise distance between colonies of a different ST isolated from the same participant was 60 SNVs. This was between ST59 and ST778 on participant F5. The minimum pairwise

difference between participants was 151 SNVs and was between ST59 on participant M7 and ST778 on participant F5. Therefore, this is likely to be a recent divergence rather than two distinct, well-developed lineages. The maximum pairwise distance between sequence types was 15,384 SNVs suggesting the presence of different lineages that have substantially diverged over time. The median pairwise distance was greater within participants (4517 SNVs) than between participants (3859 SNVs) suggesting that differing sequence type plays a greater role in diversity than different source.

3.4.6.5 Assessment of potential thresholds to define *S. epidermidis* subtypes

Pairwise distance thresholds have previously been employed to define subtypes of *S. aureus* (285,317). These data suggest five potential thresholds (Figure 3.5, Numbers 1 – 5) define subtypes of *S. epidermidis* isolates according to SNV differences.

1) The maximum pairwise distance between colonies of the same ST overall,
2) the maximum pairwise distance of colonies of the same ST isolated from the same participant,
3) the minimum pairwise distance between colonies of a different ST isolated from different participants,
4) the minimum pairwise distance between colonies of a different ST,
5) the minimum pairwise distance between colonies isolated from different participants.

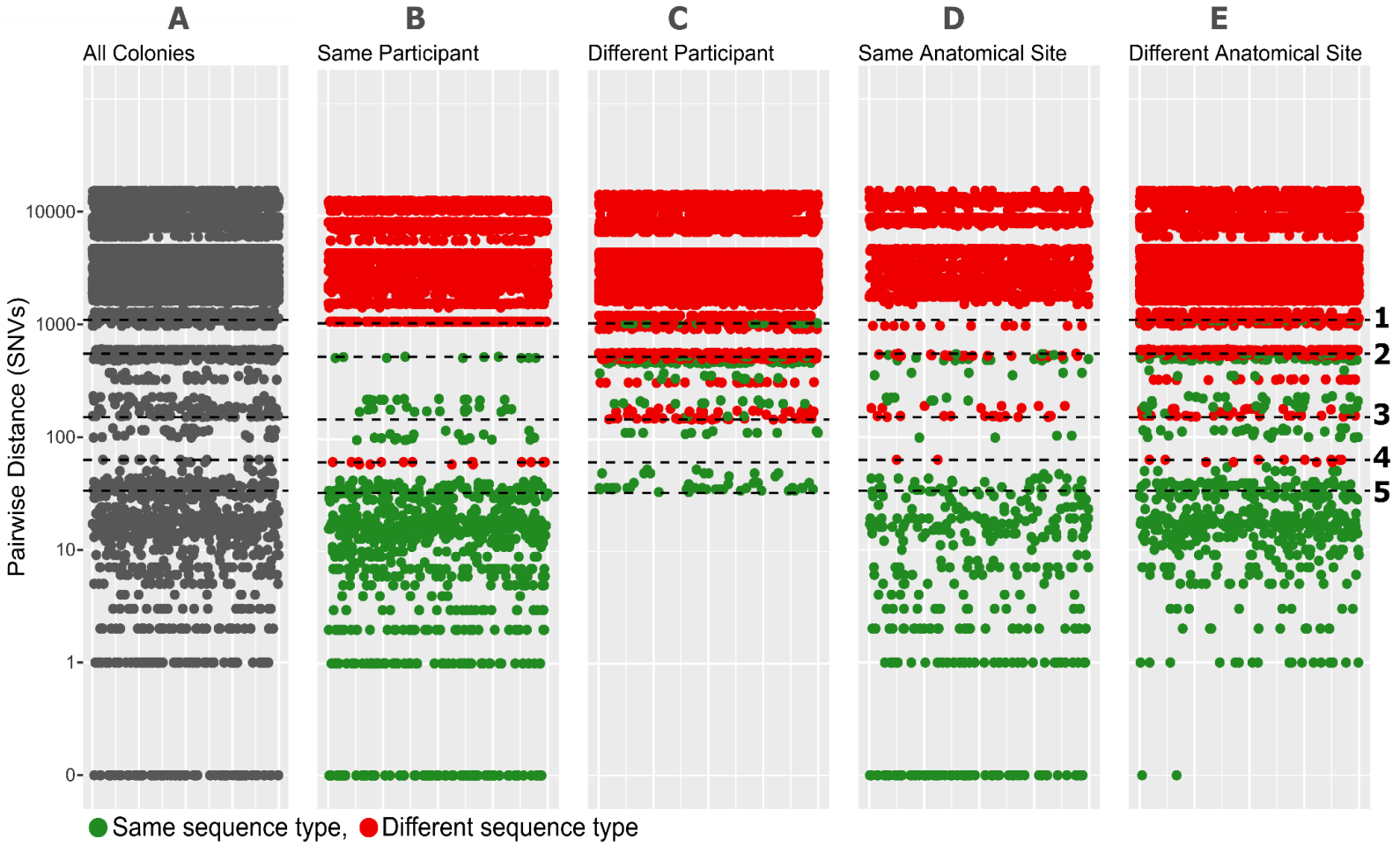


Figure 3.5. Dot plot of pairwise distance between colonies. From: A) all participants and anatomical sites, B) from the same participant, C) from different participants, D) from the same anatomical site within a participant, E) from different anatomical sites within a participant. Green dots represent colonies of the same ST, red dots represent colonies from different STs. Each dot represents the pairwise distance measured between 2 colonies. Black dashed lines show thresholds of 1) 1095 SNVs, 2) 550 SNVs, 3) 151 SNVs, 4) 60 SNVs and 5) 34 SNVs.

3.4.6.5 Pattern of diversity by sequence type

To determine whether different patterns of diversity are observed within different STs the median pairwise distances between colonies of each ST isolated from participants were plotted. Additionally, I wanted to explore the two most clinically useful thresholds.

- The maximum pairwise distance between colonies of the same ST isolated from the same participant is an indicator of the upper limit of diversity within a clone. Pairwise distances below this threshold could be used to infer recent transmission.
- The minimum pairwise distance of colonies of the same ST isolated from different participants is an indicator of diversity between clones. Pairwise distances above this threshold could be used to refute recent transmission.

Some of the STs were isolated from more than one participant and thus had more than one pairwise distance measured. These were ST59, ST73 and ST297 and were coloured blue, red and green, respectively. ST297 was isolated from two different participants at different time points therefore between-participant comparisons were not made to remove the potential of longitudinal mutations confounding the data. Furthermore, five STs (ST17, ST543, ST767, ST768) were detected only once at the same time point therefore pairwise comparisons could not be made.

27/30 (90%) of the median pairwise distances from colonies of the same ST isolated from the same participant ranged from 1 to 100 SNVs (Figure 3.6, Panel A.) and overall, the median was 13 SNVs (IQR = 1 – 18 SNVs). A similar amount of diversity is apparent within different STs.

The maximum pairwise distance of colonies of the same ST isolated from the same participant again showed a similar amount of diversity within individual STs with 24/30 (80%) of pairwise distances ranging from 1 to 100 SNVs (Figure 3.6, Panel B.). Overall, the median was 19 SNVs (IQR = 2 – 37 SNVs).

The minimum pairwise distance of colonies of the same ST isolated from different participants could only be measured for two STs and in eight instances (Figure 3.6, Panel C.). This ranged from 34 to 1095 SNVs as also seen in Figure 3.5 and the median was 434 SNVs (IQR = 319 – 640 SNVs).

Importantly the plots show that there is overlap between the maximum diversity within a participant and the minimum diversity between participants. This is markedly different to what has been observed in nasally carried *S. aureus* where populations are discrete between hosts. These different observations between *S. aureus* and *S. epidermidis* could not be accounted for by differences across anatomical sites. The pairwise distance of colonies of the same ST of *S. epidermidis* isolated from

different anatomical sites on the same participant ranged from 0 – 550 SNVs with a median of 17 SNVs (IQR = 11 – 29 SNVs).

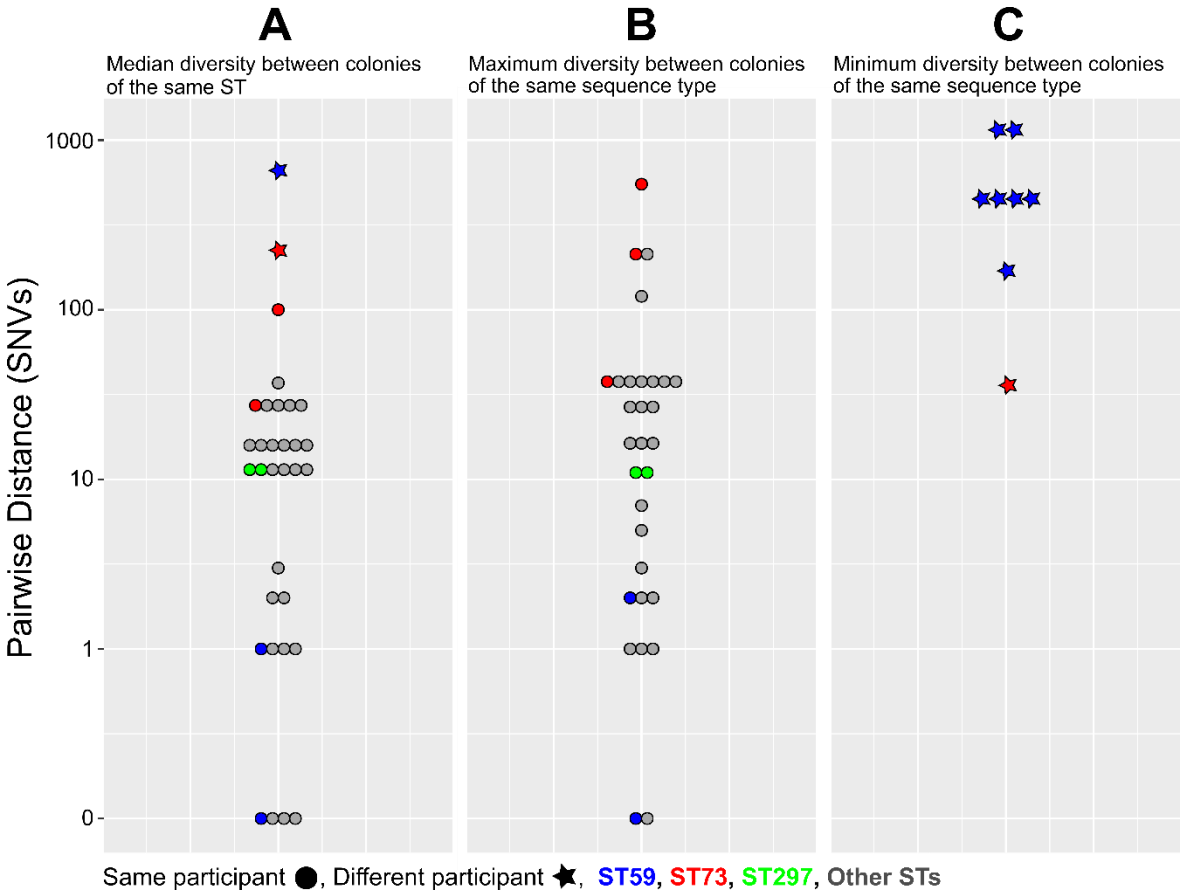


Figure 3.6. Dot plots of maximum, minimum, median and mean pairwise distances between colonies of the same ST isolated at the same time point. Pairwise distances of colonies isolated from the same participant are represented by circles and by stars for those isolated from different participants. Sequence types isolated from more than one participant are coloured.

3.4.6.7 Pattern of diversity by within-host sequence type

To assess the amount of within-host genomic diversity of *S. epidermidis* occurring within study participants, pairwise distances between two colonies of the same sequence type isolated from the same participant at the same time point were measured (Figures 3.7a – 3.7e). According to this criterion, sequence types in which only one colony was isolated from a participant at a single time point were unable to be analysed. As above this resulted in five sequence types being excluded from the analysis as only one sample was isolated. Also, as above, ST59, ST73 and ST297 were isolated from more than one person and are coloured blue, red and green, respectively. There were 938 pairwise distances in total. Pairwise distances ranged from 0 to 550 SNVs with a median of fourteen SNVs (IQR 2 – 21 SNVs). There were 50/938 instances of the pairwise distance between genomes being over 100 SNVs and these occurred in 3/30 (10%) of the within-participant STs. The most frequent pairwise distance was 0 SNVs occurring 162 times and in 24/30 of the within-participant STs. Overall, the range of pairwise SNV distance within-ST, within-participant did not vary a great deal and there were no individual STs that stood out as a major outlier.

Figure 3.7a. Dot plot of pairwise distance of SNVs between colonies of the same ST isolated from the same participant at the same time point. ST59, ST73 and ST297 are coloured blue, green and red respectively and were isolated from more than one participant.

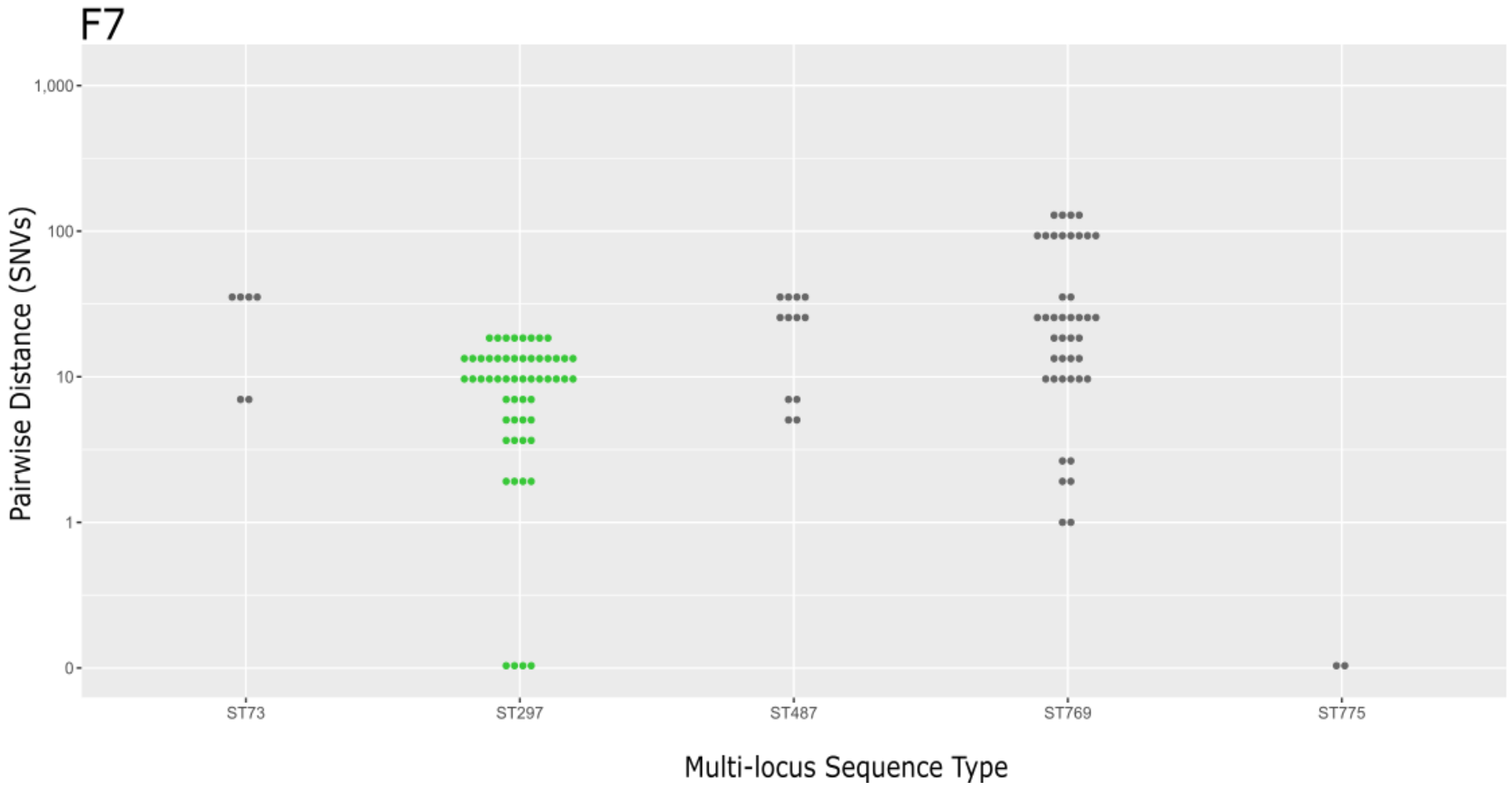


Figure 3.7b. Dot plot of pairwise distance of SNVs between colonies of the same ST isolated from the same participant at the same time point. ST59, ST73 and ST297 are coloured blue, green and red respectively and were isolated from more than one participant.

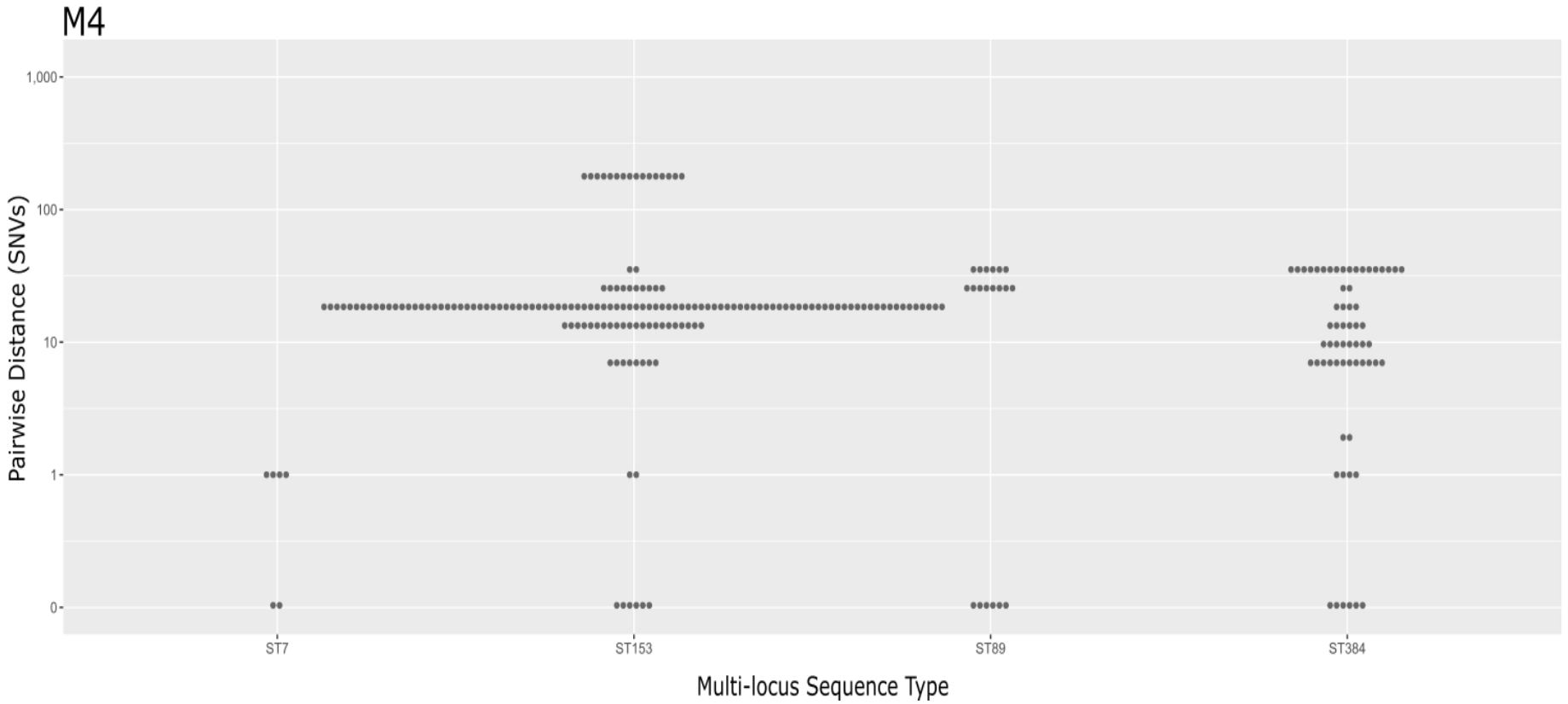


Figure 3.7c. Dot plot of pairwise distance of SNVs between colonies of the same ST isolated from the same participant at the same time point. ST59, ST73 and ST297 are coloured blue, green and red respectively and were isolated from more than one participant.

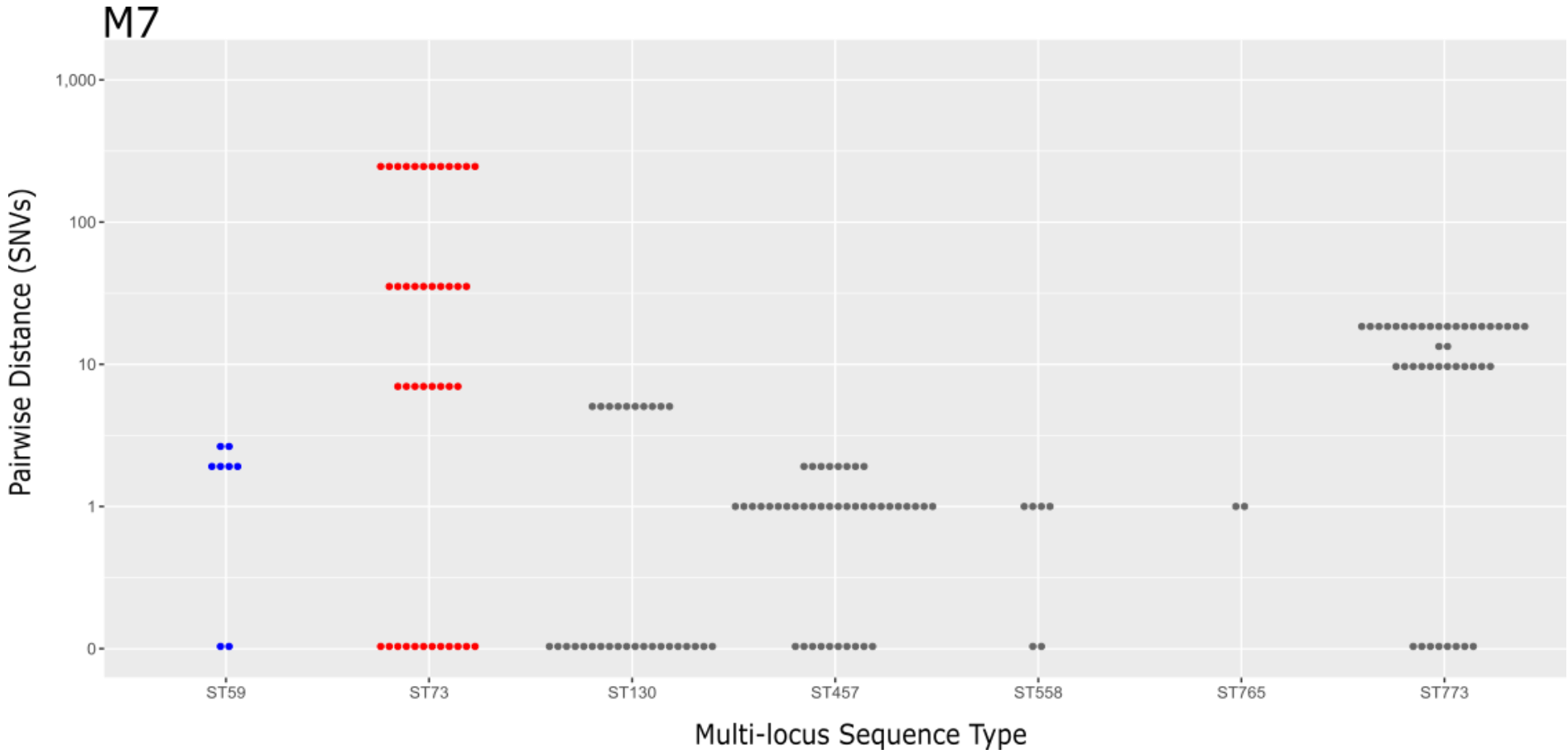


Figure 3.7d. Dot plot of pairwise distance of SNVs between colonies of the same ST isolated from the same participant at the same time point. ST59, ST73 and ST297 are coloured blue, green and red respectively and were isolated from more than one participant.

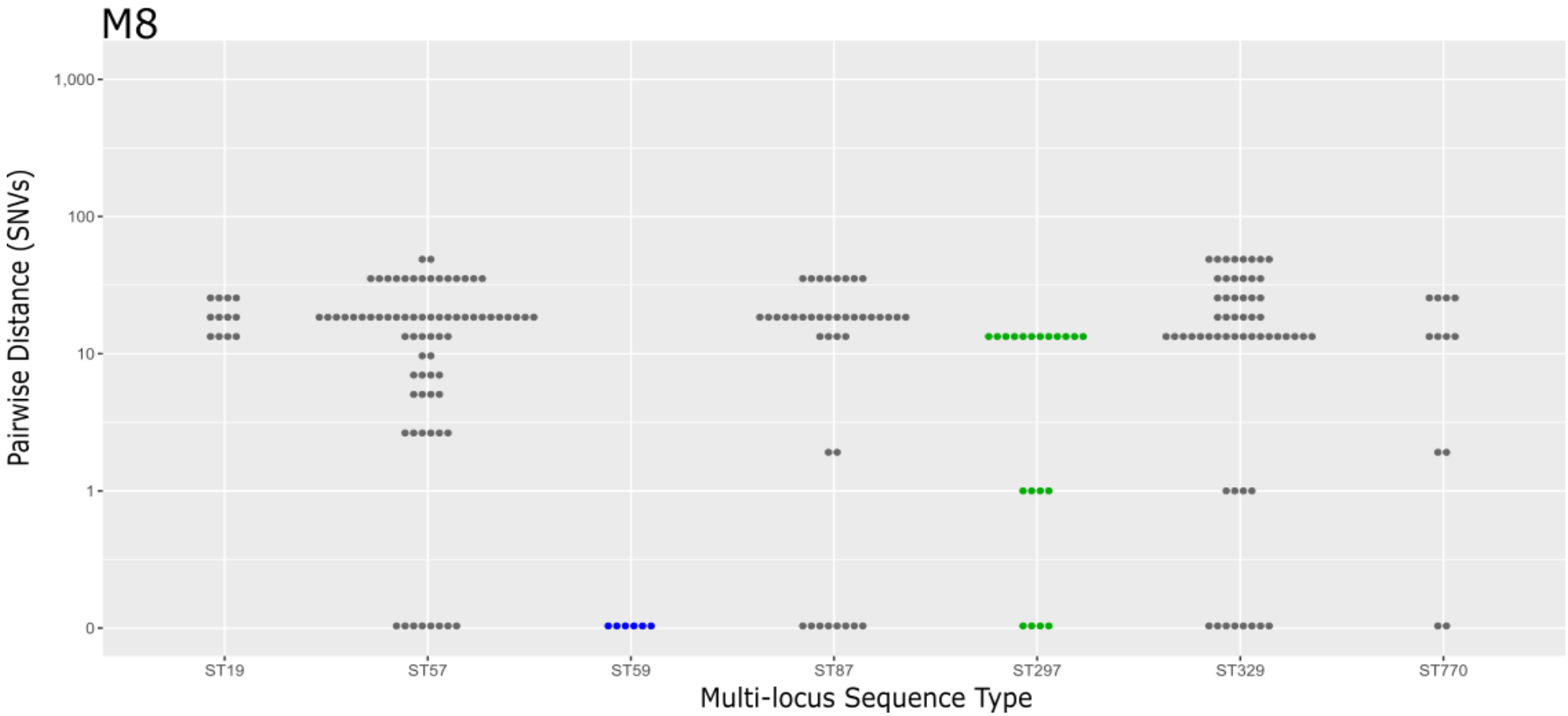


Figure 3.7e. Dot plot of pairwise distance of SNVs between colonies of the same ST isolated from the same participant at the same time point. ST59, ST73 and ST297 are coloured blue, green and red respectively and were isolated from more than one participant.

3.4.7 Calculating the Positive Predictive Value that colonies were isolated from the same participant

Figure 3.5, Panel C shows that the minimum pairwise distance of two colonies isolated from the same participant was 0 SNVs and the minimum pairwise distance of two colonies isolated from different participants was 34 SNVs. These two numbers represent the extreme in this dataset therefore a probabilistic approach was applied to determine the likelihood of two isolates being from the same person according to the number of SNVs between them. A positive predictive value (PPV) of all pairwise distances from the same time point, was measured whereby pairwise distances between colonies isolated from the same participant was considered a true positive and pairwise distances between colonies isolated from different participants was considered a false positive.

The relationship between pairwise SNV difference and PPV of two isolates being from the same participant is shown in Figure 3.8a for the full range of SNV differences and Figure 3.8b for SNV differences up to 10,000 SNVs. Given the previously described wide diversity of *S. epidermidis* between individuals it is not surprising that PPV falls rapidly as SNV threshold increases such that with a threshold of 100 SNVs gives 95% PPV for two isolates being from the same person. A threshold of 34 SNVs gives a 100% PPV for two isolates being from the same person which corresponds to the minimum pairwise distance of two colonies from the same ST and same person as illustrated in Figure 3.5, Panel C and Figure 3.6, Panel C.

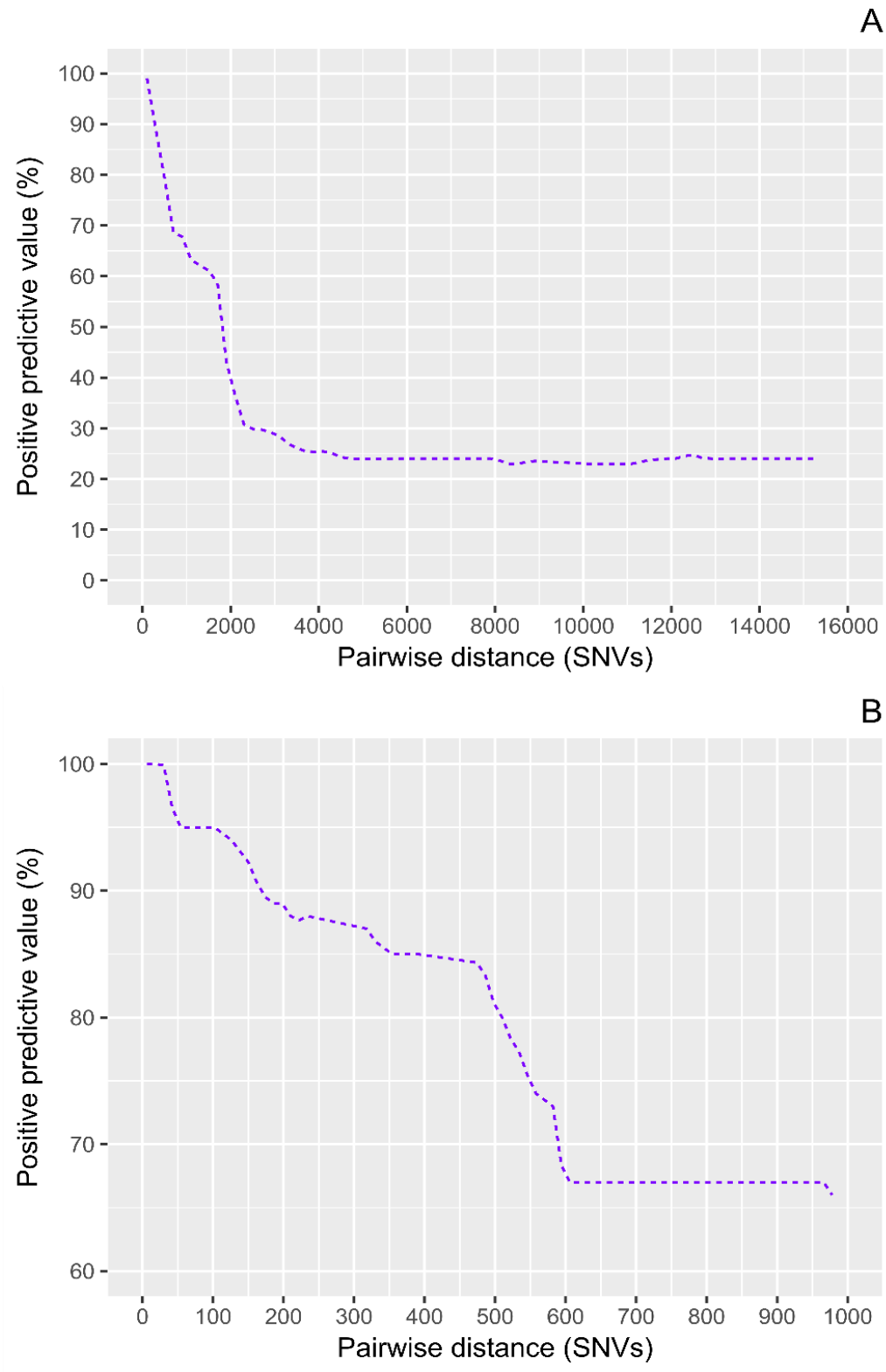


Figure 3.8. Positive predictive value of pairwise distance against probability of two colonies being isolated from the same person. A) Pairwise distance ranges from 0 – 15,384 SNVs. B) Pairwise distance ranges from 0 – 1000 SNVs.

3.4.8 Comparison of MLST, Pairwise Distance and PPV with Maximum-likelihood Analyses

Given that three different methods of describing relatedness of *S. epidermidis* isolates have been identified; MLST, pairwise distance-based subtyping and PPV, I set out to explore how to differentiate the isolates and compared them with the maximum-likelihood method of characterising clonal divergence. Maximum-likelihood trees statistically infer divergence of different lineages whereby all lineages can be traced back to a single common ancestor. Therefore, applying maximum-likelihood method to WGS data has the potential to be more accurate than MLST and measuring SNV differences alone as it not only measures differences between isolates but also shows the probabilistic branching of lineages.

3.4.9 ClonalFrameML Characterisation of Genomic Divergence of Lineages

ClonalFrameML constructs maximum-likelihood trees and shows the position in the genome where single-nucleotide polymorphisms (SNPs), homoplasy and recombination events have occurred in all branches of the tree (311). ClonalFrameML was used to construct a maximum-likelihood tree from the genomes of the 192 *S. epidermidis* carriage isolates that were mapped to reference genome ATCC 12228 with Stampy (283).

The resulting ML tree (Figure 3.9) shows a great deal of diversity between lineages. Three distinct lineages can be identified, separated by a single branch (highlighted in grey, dashed box) (Figure 3.9). One of these lineages is wholly comprised of sequence types from clonal complex 0 suggesting that the maximum-likelihood method and MLST-based phylogeny accord when characterising the population structure of *S. epidermidis*. Similarly, the other three clonal complexes represented in the ML tree are on branches separated from other isolates. The three main lineages have diverged through the accumulation of SNPs, recombination events and a great deal of homoplasy. There is more recombination within sequence types than between sequence types. Moreover, there is more recombination occurring within CC5 than without. This concurs with the conclusions of Miragaia *et al.* (2007) (238) that this clonal complex - named CC227 at the time - contains much recombination and may be the result of conservation of restriction-modification systems (238).

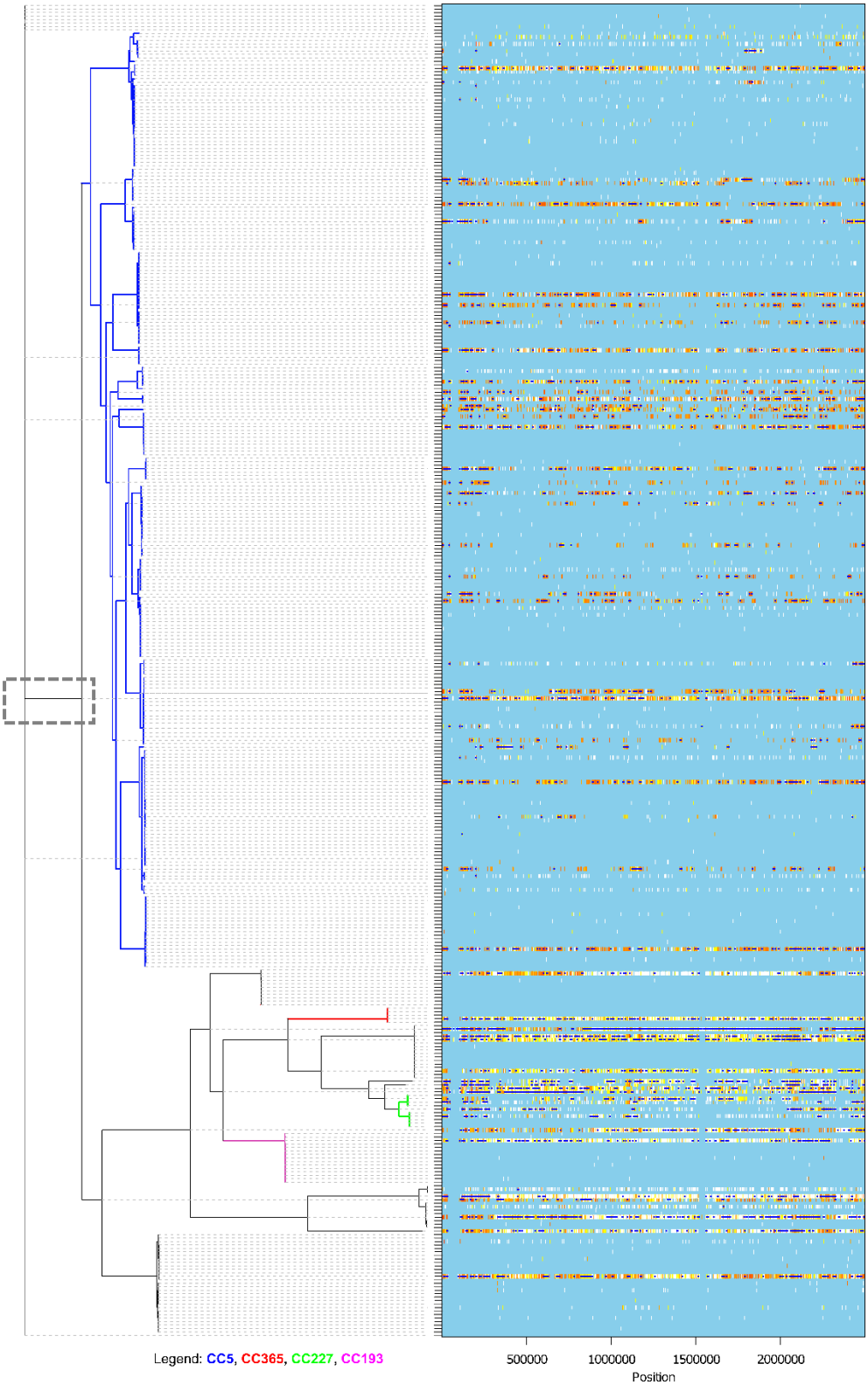


Figure 3.9. Maximum-likelihood tree of 192 carriage isolates mapped to ATCC 12228 constructed using ClonalFrameML showing tree (left) and position in the genome of polymorphisms (right). Coloured branches show branching of isolates belonging to the four largest clonal complexes according to MLST. In the genome white vertical lines indicate single nucleotide polymorphisms, yellow to orange vertical lines indicate homoplasy and blue horizontal lines indicate recombination events.

3.4.10 Comparison of Maximum-likelihood method, MLST and SNV-based subtyping to characterise *S. epidermidis* carriage

Output from the three different methods of maximum-likelihood, MLST and SNV-based subtyping was compared to assess how they characterised the relatedness of the 192 carriage isolates (Figure 3.10). The ML tree is the same as that seen in Figure 3.9. Due to the high diversity amongst the *S. epidermidis* isolates all branch lengths of the ML tree were cube-rooted using the phylogenetic analysis APE package (319) in the statistical software R (320). This reduced branches relative to their length. The tree was rooted to its mid-point using iTOL (316) under the assumption there is a single common ancestor.

Each node of the tree represented a single isolate and was aligned with both its sequence type and several SNV thresholds used to cluster subtypes. Following analysis of the SNV thresholds it was decided to apply the thresholds 151, 100, 60, 50 and 34 SNVs to subtype clustering. Thresholds 1095 and 550 SNVs were not included as was already known that they would fully concur with MLST results. 1095 SNVs represented the maximum pairwise distance of observed between two colonies isolated from different hosts. 550 SNVs represented the maximum pairwise difference observed between two colonies isolated from the same host. Instead, two other SNV thresholds were applied to the subtype clustering: 100 SNVs, the median pairwise distance giving a 95% PPV of two isolates being from the same host (Figure 3.8) and 50 SNVs, an arbitrary pairwise distance that fitted reasonably well within the other thresholds and was similar to that applied by Price *et al.* (2016) to measure transmission of *S. aureus* (317).

Genomic variation within a sequence type isolated from a single host may occur over time because of acquisition of distantly related lineage of the same ST or by polymorphic events such as recombination or hypermutation. To detect any such phenomena the time-point at which the colony was isolated was also aligned to each node on the ML tree. In addition, the participant identifiers were also aligned to tree nodes to detect lineage divergence within and between participants. The branch separating the 3 same distinct lineages from Figure 3.9 is highlighted in a grey dashed box.

At a threshold of 151 SNVs 30/32 (94%) of the STs concurred with the subtype clusters. At a 100 SNV threshold 29/32 (91%) of the STs concurred with the subtype clusters. Then at a threshold of 60 SNVs it was 28/32 (88%). At thresholds of 50 and 34 SNVs concurrence then increased back to 29/32 (91%). Discrepancies arose in which isolates of the same ST belonging to different subtypes (ST73, ST297 and ST769 in Figure 3.11, Panel B and Panel C). A discrepancy arose in which isolates of the same subtype belonged to two different STs (ST59/ST778 in Figure 3.11, Panel A). Discrepancies were associated with colonies being isolated from different hosts in 3/4 of discrepant STs and were not

associated with colonies being isolated at different time points. ST153 and ST775 differed by one SNV in the MLST loci, in the *yqiL* gene locus. They differed by 167 – 189 SNVs (median = 184 SNVs, mean = 183 SNVs) in their genomes. Isolates belonging to ST73 differed by 0 – 563 SNVs (median = 220 SNVs, mean = 252 SNVs). The discrepancies observed in ST297 and ST769 were a result of the 100 SNV and 60 SNV thresholds respectively clustering the STs into different subtypes.

ClonalFrameML was used to characterise the genomic events causing discrepancies between STs, subtype clustering and maximum-likelihood analysis. Discrepancies occurring in isolates belonging to ST73, ST59 and ST778, and ST153 and ST775 were the result of SNVs and recombination events separating isolates from the same ST or by minor changes in the MLST loci resulting in different STs that were genomically highly related.

Figure 3.10. Maximum likelihood tree of 192 genomes of *S. epidermidis* carriage isolates. Genomes were mapped to ATCC 12228 and the tree was constructed using ClonalFrameML. Branch lengths were square rooted to reduce variation in length. The tree was rooted to the mid-point. White dots represent samples isolated at time point 0, black dots represent samples isolated at time point 12. The inner ring (orange) shows the sequence type of isolates. The grey rings show clustering of subtypes according to thresholds: (from inner to outer) 151, 100, 60, 50 and 34 SNVs. The coloured identifiers on the outer-most edge denote the participant identifiers.

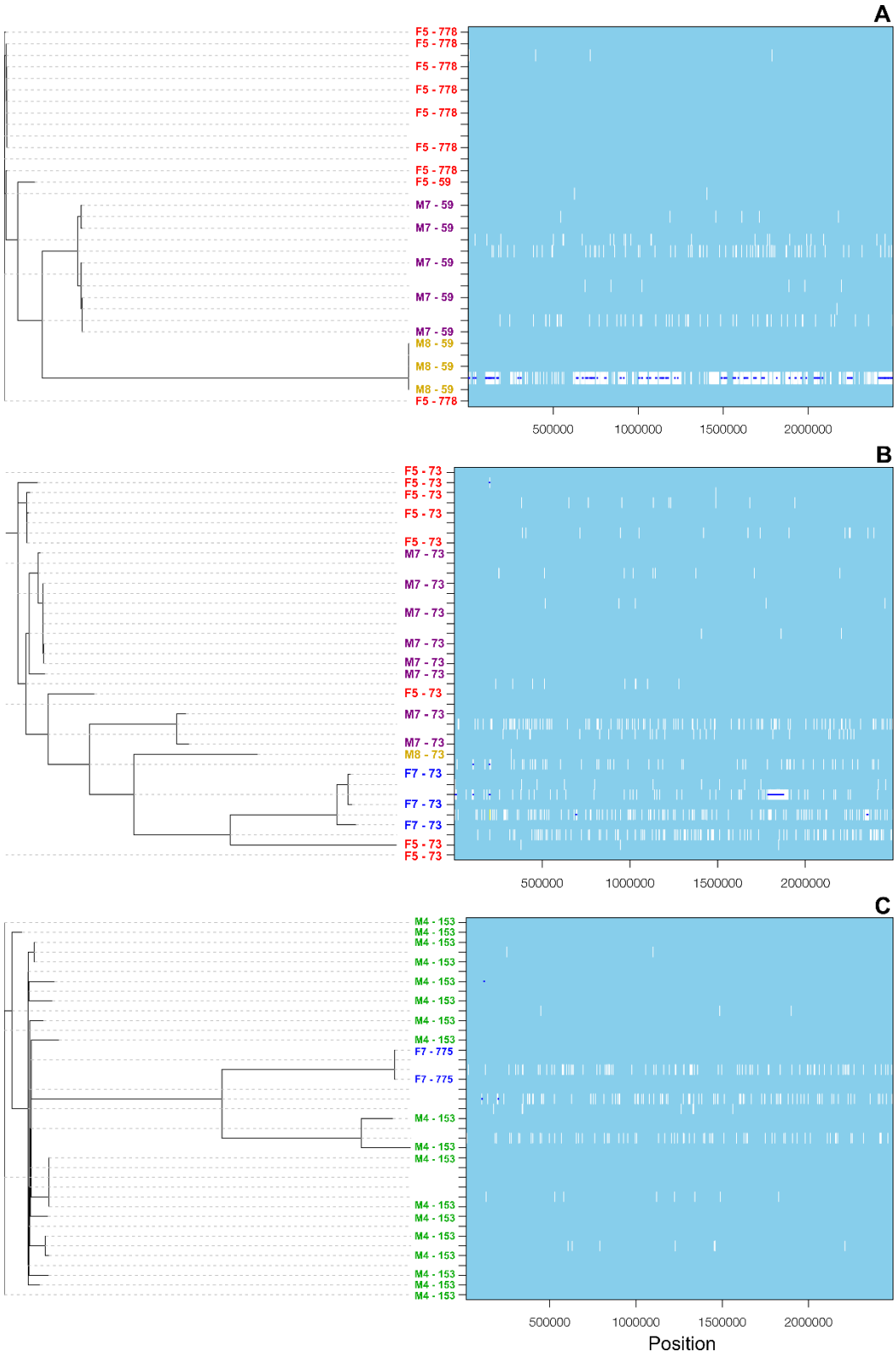


Figure 3.11. Maximum-likelihood trees drawn in ClonalFrameML observing inconsistent patterns of relatedness. Isolates of ST59 and ST778 being more closely related to each other than different isolates of ST59 (panel A). Isolates of ST73 being more closely related between hosts than within hosts (panel B). Isolates of ST775 being nested between isolates of ST153 (panel C). Participant identifiers and sequence types are labelled at nodes and coloured according to participant.

3.4.11 Determination of Fixation Indices

The fixation index of subpopulations of *S. epidermidis* from participants and subpopulations from anatomical sites on participants was calculated to determine the level of recombination occurring. When looking at the indices of fixation between participants (Table 3.10) at a threshold of 550 SNVs, the most diversity observed within a ST, the fixation index is much higher. This means that if comparing only within STs most of the diversity seen at this scale may be attributed to differences between participants. With no threshold applied the fixation indices in subpopulations from participants was low indicating that the participants harboured a great deal of the diversity across the species of *S. epidermidis*. Similarly, if a threshold of greater than 550 SNVs is applied a high level of diversity was observed in a participant which can be accounted for by divergence between STs.

When looking at the fixation indices between anatomical sites on the same participant (Table 3.11) at a minimum threshold of 551 SNVs the fixation index is determined to be zero suggesting diversity within participants was dominated by differences between different STs. If a threshold of 550 SNVs or no threshold is applied the fixation index is closer to one suggesting a more clonally structured populations within anatomical sites.

Participant	Fixation Index		
	≤550 SNVs	No Threshold	>550 SNVs
F5	0.78	0.06	-0.21
F7	0.95	0.25	0.12
M4	0.94	0.54	0.33
M7	0.91	-0.12	-0.3
M8	0.97	0.05	-0.19
Mean	**0.91**	**0.16**	**-0.05**
SD	**0.07**	**0.22**	**0.24**

Table 3.10. Fixation Indices of populations of *S. epidermidis* within participants. Minus numbers are counted as 0.

Participant	Fixation Index		
	≤550 SNVs	No Threshold	>550 SNVs
F5	0.78	0.82	-2.19
F7	0.44	0.35	-0.05
M4	0.4	0.6	-0.13
M7	0.29	0.45	-0.12
M8	0.56	0.67	0.33
Mean	**0.49**	**0.58**	**-0.43**
SD	**0.17**	**0.16**	**0.90**

Table 3.11. Fixation Indices of populations of *S. epidermidis* within anatomical sites within participants. Minus numbers are counted as zero.

3.4.11 Estimating a Molecular Clock

The molecular clock of an organism is the rate at which, under normal conditions, synonymous mutations arise and become fixed in a population and are thereby observable (321,322). The molecular clock can be used to infer the date of divergence of different species by estimating the date at which the most recent common ancestor existed (321–323). This is known as the time to most recent common ancestor (tMRCA) (324). The same principle can used to infer the time of divergence of different strains (323). This method has been used to provide phylogenetic and epidemiological data on pathogen organisms (232,311,323). However, it is important to note that an organism that is ubiquitous may reside in several different "normal conditions" and therefore a single molecular clock for a species or strain may not exist.

An attempt was made to estimate the molecular clock of the strains of *S. epidermidis* isolated from the five healthy study participants using the package BactDating (325). This was done using the maximum likelihood tree discussed in Section 3.4.9 (Figure 3.9) since this phylogeny accounted for bacterial recombination (325). Figure 3.12 infers a time-scaled phylogeny of the isolates with a tMRCA as the year 1387 (CI 95% 760 – 1717). The estimated mean substitution rate was 8.07 (CI 95% 4.7 – 12.3) per year and the standard deviation of the rate of substitutions per branch was 7.59 (CI 95% 3.98 – 12.2). Figure 3.13 is a root-to-tip analysis which infers that most substitutions occurred more recently and are clusters at the lateral edge of the time-scaled phylogeny. Finally, 1000 iterations of Markov-Chain Monte Carlo methods shown in Figure 3.14. showed disparity between the mean of the parameters and the variance of the parameters, therefore there was no convergence between the methods. Overall, the large CI 95% of the tMRCA and the results of the Markov-Chain Monte Carlo traces indicate that the analysis was not successful and was not a reliable inference.

A second attempt to infer a molecular clock was made using a bespoke method developed by Professor Daniel Wilson whereby isolates were separated according to the timepoint of isolation (month 0 and month 12) and assigned to subclusters of 170 SNVs or less. This generated 31 subclusters. Figure 3.15. shows that in each subcluster the number of SNVs between isolates from the two timepoints was plotted against time (in days) as a form of root-to-tip analysis. Most promising was Cluster 4 which showed an increase in the number of SNVs between isolates from month 0 (on the left-hand of the *x*-axis) and month 12 (on the right-hand of the *x*-axis). However, the small sampling frame for this subcluster alongside the incongruous results in the other subclusters meant this method also did not generate a reliable inference of the molecular clock. A larger collection of isolates more clonally related (with smaller pairwise distances) from timepoints

0 and 12 would have generated a larger subcluster from which a more reliable inference may have been derived.

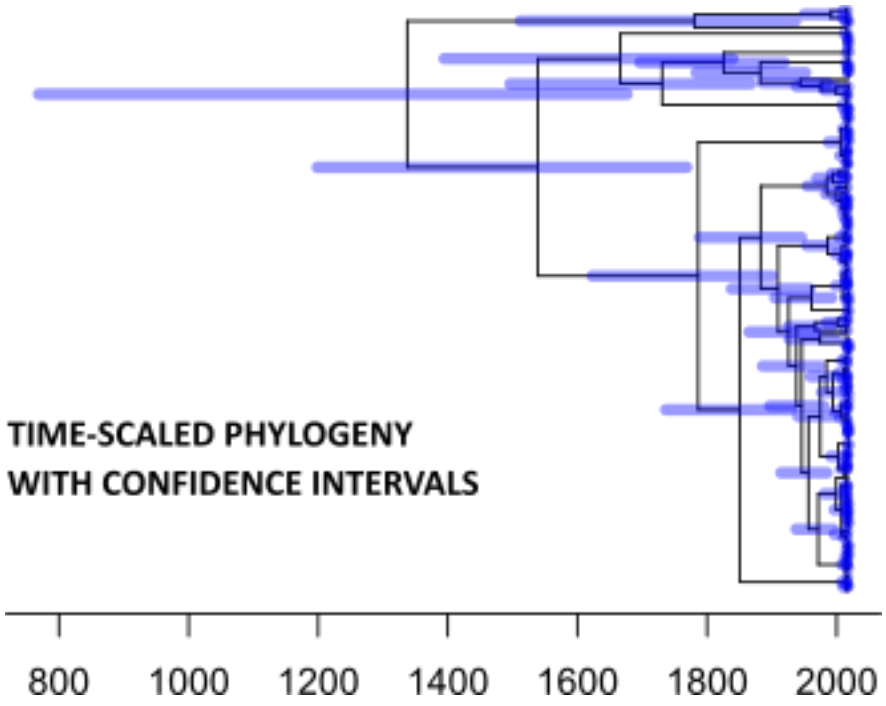


Figure 3.12. A time-scaled phylogeny of the 192 colonies of *S. epidermidis* isolated at timepoint 0 and timepoint 12. The inferred time of the existence of the most recent common ancestor was 1387 however the CI 95% dated from 760 to 1717.

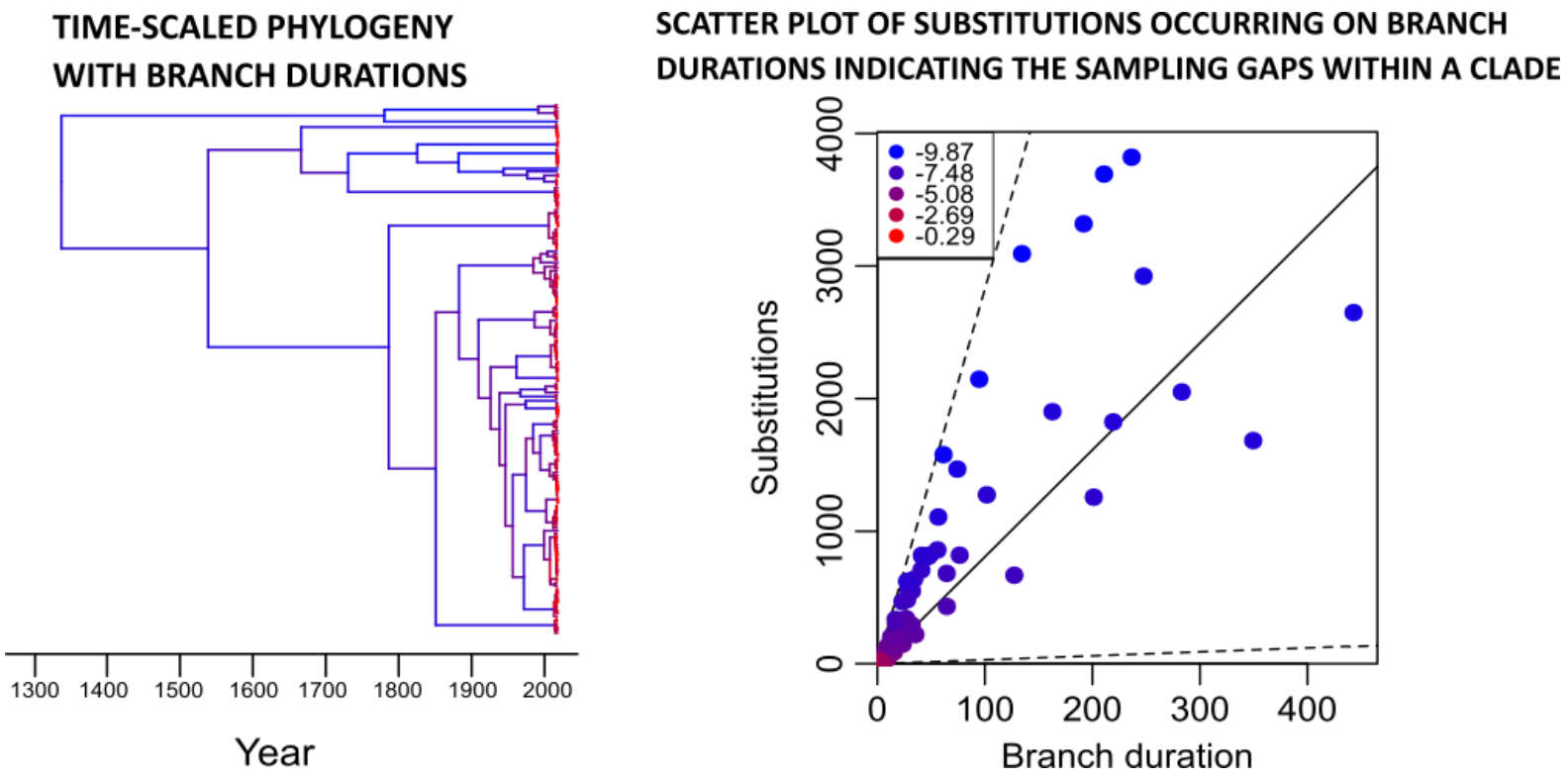


Figure 3.13. A root-to-tip analysis of the 192 colonies of *S. epidermidis* isolated at timepoint 0 and timepoint 12. The plot shows most substitutions occurred most recently.

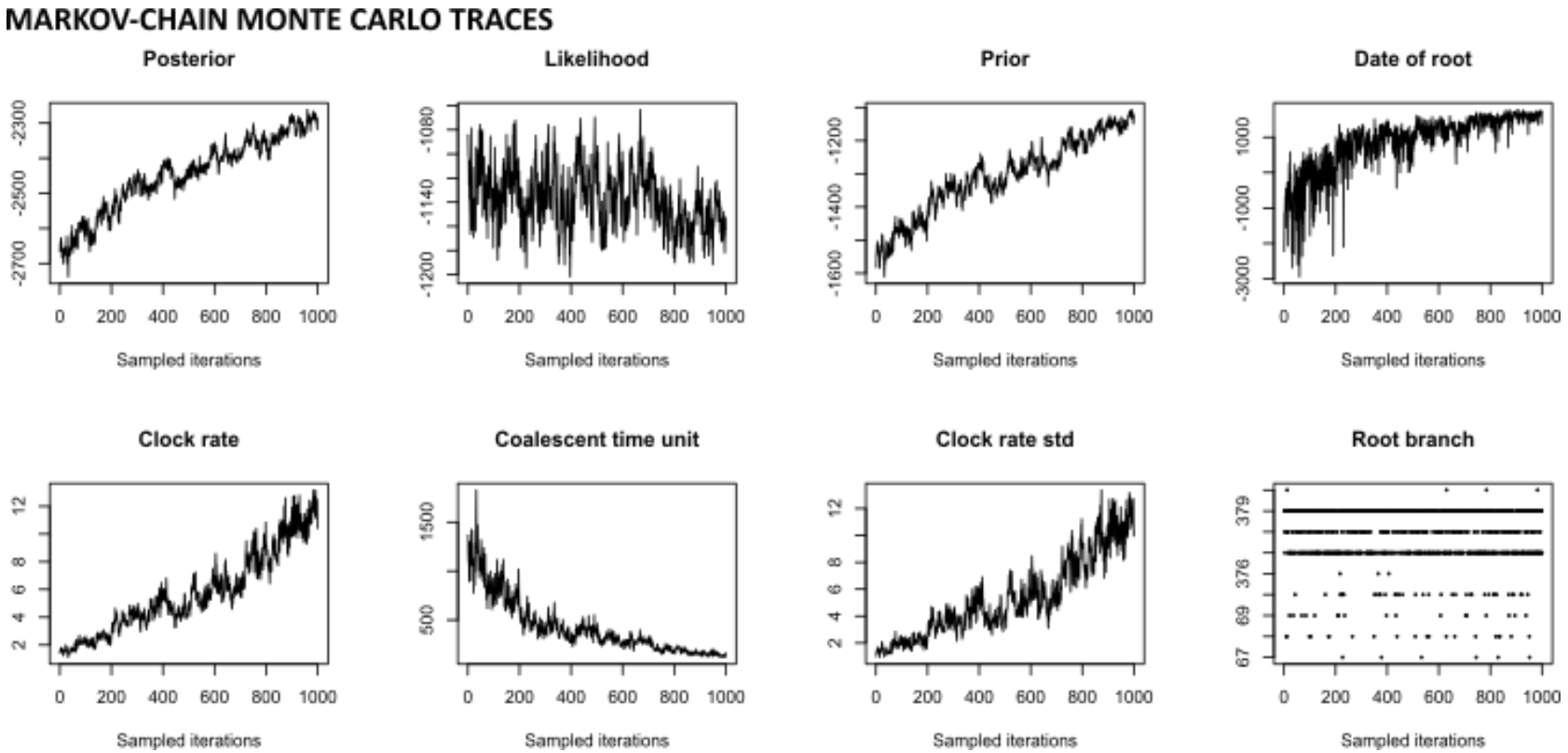


Figure 3.14. Output from Markov-Chain Monte Carlo methods using BactDating. A lack of concordance between the means and variance of the methods indicates a lack of convergence

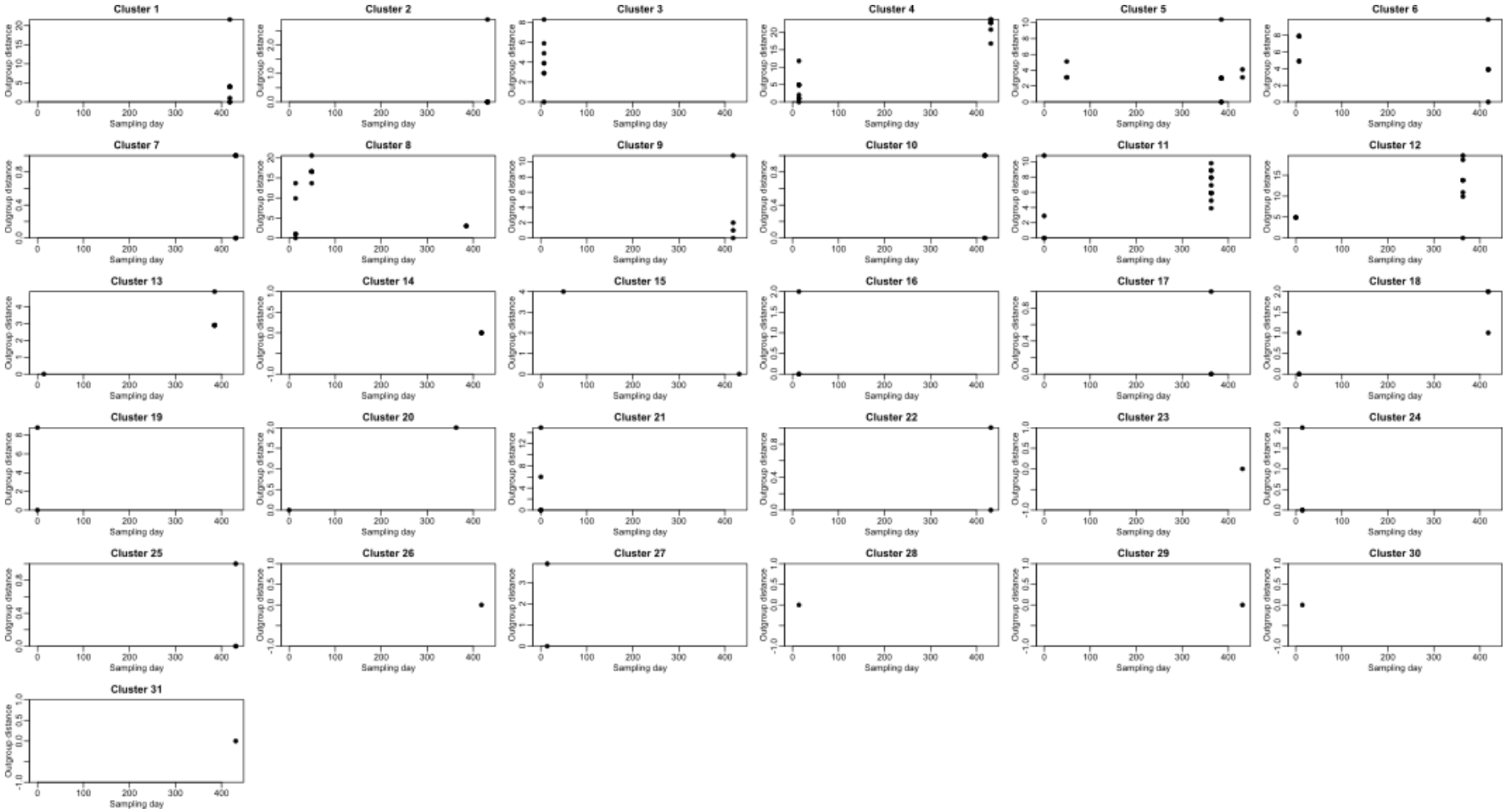


Figure 3.15. Thirty-one subclusters of the 192 colonies of *S. epidermidis* isolated at timepoint 0 and timepoint 12. Isolates related by 170 SNVs or less were subclustered. The number of SNVs (outgroup distance) was plotted against the number of days between sampling times. The marked diversity of the isolates resulted in the high number of subclusters from which a reliable inference of molecular clock could not be derived.

3.5 Discussion

Thomas, Miragaia *et al.* (2007) developed the current MLST scheme (246) and used it to characterise the phylogeny of *S. epidermidis* (238). From the 217 isolates they investigated they detected 74 different sequence types of which 74% belonged to a single dominant clonal complex, CC2 using goeBURST (234,238). The isolates investigated were of a clinical origin (238) and there has been little work exploring the population structure of *S. epidermidis* in healthy human carriage. The same group suggested applying a combination of pulsed-field gel electrophoresis (PFGE) and staphylococcal chromosome cassette *mec* (SCC*mec*) typing as a discriminatory method to give information on both contemporaneous epidemiology and long-term clonal evolution (236). Whilst this method may be applicable to clinical isolates where up to 77% carry *mecA* (139) far fewer commensal isolates (1 - 11%) harbour the gene (98,99,105). More recently, the use of metagenomics has shown that subtypes of *S. epidermidis* are generally conserved within individuals and are distributed across anatomical sites at varying relative abundance (260)The authors found that the feet carried a markedly lower diversity of subtypes than sebaceous sites where a greater diversity was observed though with some strains still dominating (260). Furthermore, temporal stability of population subtypes within individuals and anatomical sites has been demonstrated (261). The work in this thesis chapter set out to explore the diversity of 192 commensal isolates of *S. epidermidis* according to several different methods; multi-locus sequence typing (derived from whole-genome sequencing data), and three approaches using whole-genome sequencing data; SNV-based subtyping, SNV-based positive predictive value and maximum-likelihood methods. In this study 32 different sequence types were detected. Twelve were previously undescribed and comprised both novel profiles of known alleles and novel alleles. The novel STs belonged to several different clonal complexes suggesting that there are many unknown STs amongst the human population. 9/20 of the previously known STs have been previously associated with commensal origins with 6/9 being associated with commensalism in more than one study (45,99,120,128). According to the goeBURST method, 18/32 (56.3%) STs detected during this study belonged to the dominant clonal complex CC5 though overall, the 32 STs were spread across the known MLST phylogeny of *S. epidermidis*. In-keeping with current literature, the goeBURST-based minimum spanning tree showed there was a single dominant clonal complex CC5, previously known as CC2 which is currently comprised of 466/782 (59.6%) of the STs including some associated with healthy human carriage and has ST5 as the proposed founder (117,120,238). It has been suggested that dominance of this clonal complex is due to an unusually high rate of recombination giving rise to a greater number of STs than other clonal complexes (117). Whereas the work of Miragaia *et*

al. (2007) (238) showed that recombination is a strong driver of diversity amongst clonal strains of *S. epidermidis*, the maximum likelihood tree constructed using ClonalFrameML in this study shows recombination also occurring a great deal amongst commensal community isolates outside of the dominant clonal complex.

As was to be expected, my analysis measuring pairwise SNV distances between isolate genomes showed that there is much greater diversity between STs than within STs. The presence of multiple different STs in each participant therefore had a strong impact on the overall measured genomic diversity of *S. epidermidis* carriage. When isolated from the same participant the maximum within-ST pairwise distance was 550 SNVs. This was markedly higher than the 26 SNVs recorded in nasally-carried *S. aureus* by Golubchick *et al.* (2013) (285). The greater diversity in *S. epidermidis* was not due to sampling different anatomical sites since the pairwise distance between colonies from the same anatomical site also ranged up to 550 SNVs and had a similar median and interquartile range. This greater diversity may be due to several factors such as the ubiquity of *S. epidermidis* and the longer duration of carriage compared to *S. aureus*. When isolated from different participants the maximum pairwise distance of isolates of the same ST was 1095 SNVs suggesting that clonal expansion is a feature of transmission. SNV-based analysis has previously identified clonal expansion within STs of clinically important lineages of *S. epidermidis* (182) as well as *Staphylococcus haemolyticus* (326). The median pairwise distance of STs ranged from 0 to 100 SNVs with 15/30 being between 10 and 100 SNVs suggesting that generally there is no great difference in rates of divergence according to ST. The minimum pairwise distance between STs was 60 SNVs between the entire genomes and one SNV between the seven MLST loci. Figure 3.10 and Figure 3.11 show that genomically-related isolates may belong to different STs as a result of small amounts of recombination or the accumulation of SNVs. This has similarly been demonstrated in *Enterococcus faecium* where recombination caused both false-positive and false-negative clustering of STs (327) suggesting that MLST does not always form an accurate description of an organism's population structure. It is of note that the two of *S. epidermidis* STs being highly genomically related both involved a previously known ST and a novel ST. This could be the result of a constant process of novel STs arising from existing STs because of a small number of SNVs, some of which happen to occur in the loci used for the current MLST scheme. Conversely this may reflect a much larger population structure of *S. epidermidis* carried by humans comprised of many STs that are currently undetected. This confirms the conclusions of other studies that MLST cannot be used alone to fully define relatedness (120,236).

Previous studies differentiating isolates according to the number of SNVs across the genomes have found this a more useful method than strain-typing to fully characterise carriage, transmission and evolution of microbial pathogens (182,317,327,328). Subtyping of *S. aureus* based on pairwise distance of SNVs across the genome has been applied by Price *et al*. (2014) (328) who used a threshold of 40 SNVs previously found to be an upper limit of nasal carriage diversity (285). In this study five thresholds were selected according to conclusions of previous literature and the SNV-based analyses from this study.

These were:

1) 151 SNVs: the minimum pairwise distance between different STs isolated from different participants,

2) 100 SNVs: the median pairwise distance that gave a 95% positive predictive value that two isolates were not from the same participant,

3) 60 SNVs: the minimum pairwise distance between different STs isolated from the same participant

4) 50 SNVs: a pragmatically chosen threshold based on values from this study and cited literature examining the genomic diversity of *S. aureus*,

5) 34 SNVs: the minimum pairwise distance within STs isolated from different participants.

A threshold pairwise distance up to 100 SNVs gave a 95% positive predictive value that isolates of *S. epidermidis* originated from the same participant in this study. This value could be used to suggest a recent transmission event if isolates from different people were within 100 SNVs however more carriage isolates from more participants would need to be analysed to better represent *S. epidermidis* in the human population. While the threshold thus identified may be applicable to other settings, the specific posterior predictive value would strongly depend on the sampling details. For example, isolates from protracted-stay in-patients may have a significantly reduced diversity of *S. epidermidis* and so the number of SNVs generating a 95% PPV would be reduced Different lineages characterised in the maximum-likelihood trees constructed in ClonalFrameML accorded with the subtypes derived from the 100 SNV threshold. Figure 3.10 shows that isolates of ST297 isolated from different participants formed discrete subtypes up to 100 SNVs and that they also existed on different branches of the maximum-likelihood tree meaning whilst they belonged to the same ST there was clonal divergence. Figure 3.10 also shows that branches separating isolates on the maximum-likelihood tree concorded

with subtypes based on thresholds of 60, 50 and 34 SNVs. Subtyping carriage isolates based on the number of SNVs that separate them may be used to confirm if they are related and thus either from the same person, same site of infection or involved in transmission. Although this study has concentrated on commensal carriage the same methods could potentially be applied to inform on epidemic or persistent nosocomial strains of *S. epidermidis*. Price *et al.* (2014) used a threshold of 40 SNVs to characterise *S. aureus* transmission in an intensive care unit and found that acquisition was rarely a result of person-to-person transmission suggesting an undetected source (328). Furthermore, work has already been done investigating the number of SNVs that separate colonising populations and infecting populations of *S. aureus* (318). Identifying genomic relatedness between infecting and carriage populations of *S. epidermidis* could aid infection prevention and management.

The diversity observed in this study is in accordance with that seen in other studies (99,120,261,329). Oh *et al.* (2016) used metagenomic sequencing to identify 14 different subtypes of *S. epidermidis* carried by healthy individuals. They found that diversity was both higher and more temporally stable than observed during this study (261). This suggests that if more colonies were analysed in this study, then a greater diversity of STs would be detected at the two time points and that the same STs would be detected at the two time points more frequently.

3.6 Conclusion

The work undertaken in this chapter has demonstrated that healthy participants carry multiple sequence types of *S. epidermidis* both within and between multiple anatomical sites.

The frequency at which I identified novel sequence types spread right across the known phylogeny of *S. epidermidis* according to MLST indicates substantial, currently unmeasured diversity in the population of this species.

Diversity between different STs represented as little as 0.002% of the genome with single SNVs between the alleles used for MLST causing highly related isolates to be assigned to different STs.

Isolation from different participant played a small but still significant role in the relatedness of isolates of the same ST.

Measuring pairwise distance between isolate genomes is thus a more reliable method of determining relatedness than sequence typing. This can be used to determine a probability of same origin or to subtype isolates as has been documented in *S. aureus* (328,330).

The inability to reliably infer a tMRCA and molecular clock was largely due to the high diversity of the isolates which is indicated by the diversity of the multi-locus sequence types identified. Analyses more focused on a greater number of isolates of the same ST and isolated from more timepoints would perhaps produce a more reliable inference.

My findings indicate thresholds for relatedness of *S. epidermidis* isolates which are different from those used previously for *S. aureus* and should be evaluated in clinical studies, for example, assessing transmission in healthcare settings.

WGS has been used to inform the evolution and population dynamics of clinically significant *S. epidermidis* (182,331). Thus far, WGS analyses of the commensal *S. epidermidis* has been very limited. This study addresses that issue by investigating healthy carriage of *S. epidermidis* which comprises our main interaction with the organism.

Chapter 4

Using whole-genome sequencing data to discriminate between *Staphylococcus epidermidis* orthopaedic device-related infection and contamination

4.1 Introduction

Advances in surgery and infection prevention during orthopaedic device implantation procedures mean the proportion of people developing an associated infection are low at 1 – 2% (35,332,333). However, the increased number of implants being performed every year has led to an overall increase in the incidence of orthopaedic device-related infection (35,333). Although less common than aseptic complications, ODRI can be devastating with serious morbidity and cost implications (35,334–337).

Diagnosis of ODRI may rely on clinical, haematological, histological, and microbiological findings. For microbiological analysis, it is recommended that five or six specimens of periprosthetic tissue are taken intra-operatively, each with a different set of sterile instruments (61,273). Current UK guidelines recommend that specimens are aseptically inoculated either to nutrient-rich agar and broth for direct culture or into blood culture bottles for automated detection of microbial growth (60). Culturing the sonicate of an explanted prosthesis may increase sensitivity, particularly for organisms in a biofilm (86,273,338) however this comes with a greater risk of contamination (62) and should not replace the analysis of periprosthetic tissue (60). The culture of indistinguishable organisms from more than one specimen from a set of samples is one of the diagnostic criteria used to help to identify the cause of infection (76).

Staphylococcus epidermidis is one of the organisms most commonly associated with ODRIs (43,47,54,339). However, it is also a common cause of specimen contamination meaning the culture of *S. epidermidis* does not necessarily indicate true infection. It may be difficult, when using conventional methods, to demonstrate that isolates of *S. epidermidis* from multiple samples are of the same clone. Using the antimicrobial susceptibility profile as a surrogate marker of relatedness is unreliable since resistance profiles that are common to unrelated isolates exist, particularly regarding nosocomial strains. Similarly, MLST does not reliably indicate that isolates are identical due to the prevalence of particular sequence types in hospital settings. Numerous studies have attempted to determine if particular strains or genes are more indicative of infection or contamination however, they have thus far been unable to discriminate between nosocomial infection and nosocomial carriage (109,111,238,255,256).

Recent studies have assessed the potential for clinical metagenomics to detect and identify organisms from bone and joint specimens (340–342). Two studies applying metagenomic sequencing to synovial fluid showed promising sensitivity (94.9% and 90% identifying the genus and 86.5% and 83% identifying the species) when compared with culture of synovial fluid and intra-operative specimens suggesting pre-operative specimens are also suitable for this method

(343,344). Another study applied metagenomic sequencing to culture-negative specimens from cases of suspected PJI (345). The authors found that targeted antimicrobial therapy based on the sequencing results was associated with no need for subsequent intervention, however similar outcomes were observed when empirical antibiotics were administered (345). These studies have demonstrated that genome sequencing directly from the specimen might be used to supplement or potentially replace culture-based methods but such approaches are not currently suitable for routine use.

Whole-genome sequencing of cultured organisms may have the potential to bridge the gap that currently exists between culture-based phenotypic methods and metagenomic sequencing of the specimen by investigating individual colonies of cultured organisms and generating more comprehensive data than metagenomics can currently yield.

Previous studies have shown that *S. epidermidis* and *S. aureus* invasive infection is highly clonal whereby, following inoculation, the genome undergoes incremental random changes, some of which may enable adaption to the new ecological niche or to resist antimicrobial therapy (175,318,331). Different colonies may have the same MLST, PFGE or antibiogram profile but represent a heterogeneous population in the host (346). Whole-genome sequencing can be employed to track genomic changes across time (from sequential sampling) or space (from sequencing multiple colonies from the same specimen). Sequencing multiple colonies from the same specimen can be used to observe the within-host diversity of an organism at a single point in time. Multiple studies have shown that within-host diversity is dynamic and increases measurably over time (346). Whilst whole-genome sequencing would be a more time-consuming and costly method of diagnosis and analysis of ODRI in comparison to traditional methods, the information about the infection may be much more valuable. If WGS proved to be more sensitive and specific than culture at diagnosis and yielded more information about the progression of the infection, then this could have more benefits to the patient such as improved use of antibiotics. Additionally, better diagnoses may lead to reduced surgical costs which are proportionally much higher than the cost of microbiological diagnosis. Aseptic revisions cost an average of £11,897 in the UK in 2012 but septic revisions cost £21,937 (347). An improved or quicker diagnosis could potentially reduce this gap, especially if WGS was performed pre-operatively from organisms isolated from synovial fluid.

4.2 Objectives

The work presented in this chapter was aimed at developing a method of using whole-genome sequencing to discriminate between *S. epidermidis* orthopaedic device-related infection and specimen contamination. There were three aims:

1) Develop a collection of specimens of periprosthetic material that have been defined as relating to infected or to non-infected cases, according to the Musculo-Skeletal Infection Society criteria (340),
2) Determine if WGS-based population analysis methods can be used to discriminate between bacteria cultured from infected and non-infected patients (i.e., between presumptive pathogens and sample contaminants),
3) Ascertain if there is a direct relationship between the genomic diversity of the infecting population and the time from prosthesis inoculation to isolation of the organism to determine if a molecular clock of infection is observable.

Colonies of *S. epidermidis* cultured from implanted devices were investigated using phenotypic, genotypic and genomic protocols to determine their relatedness. The number of SNVs between colonies from the same specimen and specimen set was measured with the aim to establish if there is a relationship between genomic diversity and origin (infection or contamination) of the organism. The object of this work was to determine the potential of WGS as a method of measuring the diversity of the cultured population to aid diagnosis of *S. epidermidis* ODRI.

4.3 Methods

4.3.1 Sampling

4.3.1.1 Ethical Approval

All bacterial isolates described in this chapter were investigated as part of a service development and so were covered under Statutory Instrument Regulations 2002 No. 1438 (348). Isolates from the Oxford University Hospitals NHS Foundation Trust were obtained from the routine microbiology laboratory with no associated patient identifying data. Patient data associated with isolates from the Brighton and Sussex University Hospitals NHS Trust were anonymised therefore no formal ethical approval was required.

4.3.1.2 Laboratory Collection

Original cultures from specimens were collected between November 2016 and October 2018 from the John Radcliffe Hospital microbiology laboratory, part of the Oxford University Hospitals (OUH) Trust. The laboratory processes specimens obtained from the Nuffield Orthopaedic Centre (NOC), Oxford which undertakes specialisms in bone and joint infections amongst others. Original cultures were also collected from the Royal Sussex County Hospital (RSCH) microbiology laboratory between April 2017 and November 2018. The RSCH is part of the Brighton and Sussex University Hospitals Trust (BSUH). The Sussex Orthopaedic Treatment Centre (SOTC) located at Haywards Heath, is also part of the BSUH Trust and sends orthopaedic device-related specimens to the RSCH microbiology laboratory for culture-based analysis. In total 177 original cultures from 66 sets of specimens taken from 65 patients were collected. From the 177 cultures, 65 were selected for whole-genome sequencing that represented culture results ranging from one positive culture to five positive cultures. Cultures of *S. epidermidis* representing true infection and contamination were selected according to the medical microbiologist's application of the MSIS criteria (76).

Figure 4.1 summarises the number of cultures collected from each site. Original culture plates of Columbia Blood Agar or Chocolate Agar (Oxoid Ltd, Basingstoke, UK) were collected so that the diversity of *S. epidermidis* from the specimen was captured. Both RSCH and JRH microbiology laboratories used a MALDI-ToF MS (Bruker Daltonik GmbH, Leipzig, Germany) to speciate isolates.

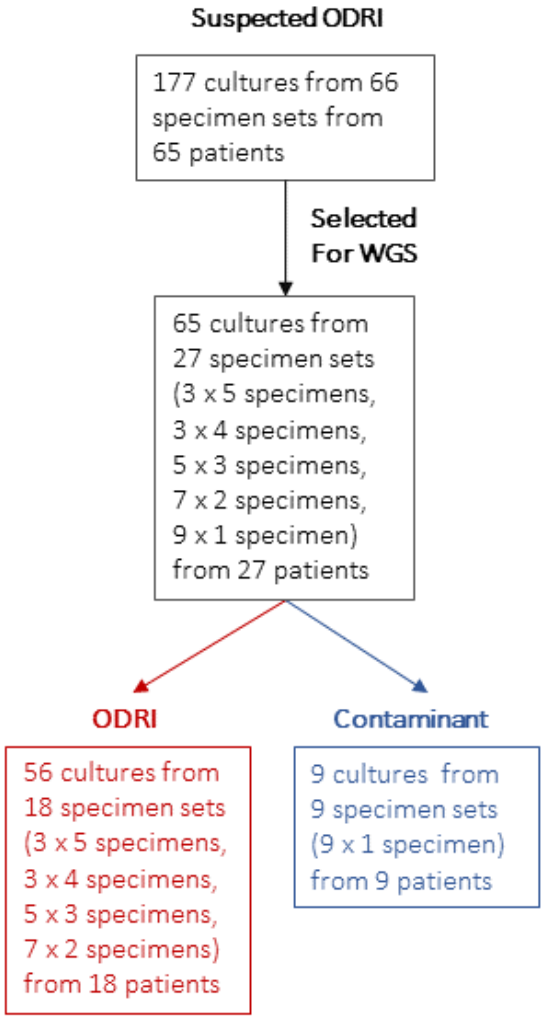


Figure 4.1. Original cultures from specimens collected and selected for whole-genome sequencing from the Royal Sussex County Hospital and John Radcliffe Hospital microbiology laboratories.

4.3.2 Determining the Clinical Significance of Isolating *S. epidermidis* from a Specimen

Diagnosis of an orthopaedic device-related infection relies on multiple investigations including microbiological culture, biochemical, haematological, histological and clinical examinations. Table 1.5 shows the new scoring system developed by Parvizi *et al.* (2018) on behalf of the Musculoskeletal Infection Society (MSIS) which aims to improve diagnostic confidence of PJI (76). Whilst the goals of treating fracture fixation infection (FFI) differs from that required for PJI many of the diagnostic criteria shown in Table 1.6 are the same and oftentimes the algorithms used for diagnosis of FFI have been adapted from those applied to PJI (265)

Therefore, for the purposes of this study, diagnostic criteria proposed by Parvizi *et al*. in 2018 (76) for infection of the hip and knee was also applied to ODRI of other anatomical regions.

Most, but not all the MSIS criteria are employed by the clinical microbiologists serving the Sussex Orthopaedic Treatment Centre and the Nuffield Orthopaedic Centre. Therefore, for the purposes of this study, where possible the MSIS scoring system was applied to establish the results of typical MSIS recommended examinations but the clinical decision by the microbiology or infectious diseases consultant was used as the ultimate determination of whether a *S. epidermidis* positive culture was the result of infection or contamination.

4.3.3 Culture and Identification

Five colonies of suspected *S. epidermidis* from each selected original culture plate were sub-cultured onto a CBA plate (Oxoid Ltd, Basingstoke, UK) and incubated overnight at 35^{o}C in aerobic conditions. Following incubation individual colonies were confirmed as *S. epidermidis* using a MALDI-ToF MS (Bruker Daltonik GmbH, Leipzig, Germany), (see Chapter 2: Materials and Methods).

4.3.4 Antimicrobial Susceptibility Testing

The Royal Sussex County Hospital microbiology laboratory tests a panel of 13 antibiotics against any staphylococcus isolated from a normally sterile tissue or fluid specimen including from medical prostheses. The John Radcliffe Hospital employs a BD Phoenix™ (Becton-Dickinson, Franklin Lakes, New Jersey, US) automated identification and susceptibility testing system to test gram-positive organisms including staphylococci. The BD Phoenix™ PMIC panel tests the MICs of multiple antimicrobial agents including cefoxitin, oxacillin, penicillin, gentamicin, rifampicin, fluoroquinolones, trimethoprim-sulfamethoxazole, erythromycin, clindamycin, linezolid, tetracycline, vancomycin and chloramphenicol. These panels are used to give the clinician a range of antimicrobials from which the most suitable can be selected. The antibiogram may also be employed as a surrogate marker of relatedness of isolates from different specimens from the same patient.

4.3.4.1 Disc Diffusion

AST was performed using the disc diffusion methods according to the British Society for Antimicrobial Chemotherapy (BSAC) (279) which are described in Chapter 2: Materials and Methods. Each sub-cultured colony of *S. epidermidis* underwent disc diffusion AST of the antibiotics listed in Table 2.1 to determine their antibiogram.

4.3.4.2 Minimum Inhibitory Concentration

The minimum inhibitory concentrations (MIC) of teicoplanin and vancomycin were tested using E-test© strips (BioMerieux, Marcy-l'Étoile, France) or M.I.C. Evaluator strips (Oxoid Ltd, Basingstoke, UK) according to the methods outlined by BSAC (279). MICs were interpreted as the lowest concentration of antibiotic that inhibited the growth of test organism as shown in Table 2.2.

4.3.5 Whole-genome Sequencing

Whole-genome sequencing was performed on an Illumina HiSeq4000 at the Wellcome Trust Centre for Human Genetics using methods described in Chapter 2: Materials and Methods.

4.3.6 Analyses

MLST

Multi-locus sequence typing was performed using a BLASTn-based method on WGS data (see Chapter 2: Materials and Methods).

Haplotrees

Haplotype trees were constructed using a maximum likelihood algorithm implemented by Professor Daniel Wilson which analyses sets of biallelic SNVs in genomes where there is no homoplasy (308). In summary short-read sequences are mapped to a reference genome to construct aligned contigs. The contigs of multiple isolates are then examined for the presence of SNVs between alleles. A table of these SNVs (or biallelic positions) is constructed to show their position in the genome. This table is then used to construct a haplotype tree which visualises how the genomes of multiple isolates differ according to the presence and location of SNVs.

Mann Whitney U Test

The Mann Whitney U Test (Wilcoxon Rank-Sum Test) is a non-parametric test between two independent groups used to determine if they likely derive from the same population (349). For this study, the Mann Whitney U test was applied to the pairwise diversity of colonies sequenced from suspected ODRI and suspected contaminant samples. A small p-value will suggest that the null hypothesis - the contaminants and infecting colonies are homogeneous and come from the same population - can be rejected (350).

4.4 Results

4.4.1 Specimens

From the microbiology laboratories of the RSCH (Table 4.1) and JRH (Table 4.2) a total of 65 specimens from 27 specimen sets from 27 patients were analysed. The MSIS criteria for defining infected or not infected and the clinical team's decision of infected or not infected was recorded.

Of the 27 patients from both hospitals the mean age was 62 years. Twelve of the 27 (44.4%) patients were female and fifteen (55.6%) were male. Nineteen of the 27 (70%) operations involved the hip, 5/27 (18.5%) involved the knee, 1/27 (3.7%) involved the clavicle in the shoulder, 1/27 (3.7%) involved the ankle and 1/27 (3.7%) involved the calcaneum of the foot heel. A set of five or six specimens was most commonly sent to the microbiology laboratory (both sites 9/27 (33.3%)) which was the recommendation of Atkins *et al*. (1998) (61). Specimens within a set growing *S. epidermidis* ranged from six being positive to a single specimen being positive.

CASE DETAILS				SPECIMEN DETAILS	
Patient	**Age**	**Sex**	**Operation**	**No. growing *S. epidermidis***	**No. for WGS**
TF12	59	F	DHS to THR	1/2	1
TF16	25	M	Review of open knee fracture wound and removal of screws	3/3	3
TF24	58	F	Revision of acetabulum	1/2	1
TF26	75	F	Washout after DHS	5/5	3
TF32	44	M	Washout of hip	1/4	1
TF56	85	M	DAIR of hip	1/6	1

Table 4.1. Table of specimens selected for analysis from the Royal Sussex County Hospital, Brighton, UK. THR: Total Hip Replacement. TKR: Total Knee Replacement. DHS: Dynamic Hip Screw. DAIR: Debridement, Antibiotics and Implant Retention.

	CASE DETAILS			SPECIMEN DETAILS	
Patient	**Age**	**Sex**	**Operation**	**No. growing *S. epidermidis***	**No. for WGS**
PDR1	73	F	One stage infected hemiarthroplasty to THR with custom assembled femoral endoprosthetic replacement	4/5	4
PDR6	67	M	First stage revision of TKR	5/6	5
PDR9	67	M	One stage revision of THR	4/6	3
PDR12	48	M	Complex acetabular reconstruction and hip replacement	5/5	5
PDR19	86	F	One stage hip revision	6/6	4
PDR17	47	M	Washout of knee and superficial wound dehiscence	1/4	1
PDR20	60	F	Exploration of hip, change of stem and acetabular liner and excision of pseudotumour	1/5	1
PDR21	46	F	DAIR of hip	5/6	5
PDR25	68	F	First stage revision of THR, complex removal of metalwork	3/7	2
PDR28	52	M	One stage revision of distal femoral replacement to total femoral replacement (Depuy Limb Preservation System endoprosthesis) with medial and lateral gastrocnemius flap 7 split skin graft	6/6	3
PDR33	76	F	Exploration hip revision and conversion to constrained liner	3/4	2
PDR36	70	F	Revision THR (gap cage/contemporary/Exeter)	4/5	2
PDR38	24	M	Treatment of acetabular metalwork infection	5/5	4
PDR39	38	M	Post anterior cruciate ligament reconstruction Removal of tibial common iliac screw, debridement and arthroscopic washout of knee joint	3/6	2
PDR44	75	M	Exploration THR, exchange of liner and head	1/5	1
PDR45	74	M	Removal of ankle screws, washout and debridement	1/5	1
PDR49	63	M	Sampling, plate removal and assessment of clavicle	1/4	1
PDR93	79	F	Revision of THR (Trabecular Metal trilogy: Trabecular Metal/augment/Exeter)	5/6	2
PDR98	79	M	Two in 1 hip revision All poly-cemented cup 200mm CPT® to trilogy with impaction of 40g osteomycin	2/7	2
PDR139	69	F	DAIR of hip for infection	3/6	3
PDR144	58	M	Calcaneum screw fixation	4/5	2

Table 4.2. Specimens selected for analysis from the John Radcliffe Hospital, Oxford, UK. THR: Total Hip Replacement. TKR: Total Knee Replacement. DHS: Dynamic Hip Screw. DAIR: Debridement, Antibiotics and Implant Retention.

Not all the analyses recommended in the new scoring system for the diagnosis of PJI by the MSIS were employed by the Royal Sussex County Hospital or the John Radcliffe Hospital laboratories to examine the cases of suspected ODRI for this study. Neither centre examined serum for the presence of D-dimer. Similarly, neither examined synovial fluid for the presence of Alpha-defensins, elevated polymorphonuclear leukocytes (PMN), C-reactive protein (CRP), white blood cells (WBCs) or leukocyte esterase therefore these tests were not included in Table 4.3.

Culturing an organism indistinguishable by antibiogram from two or more specimens from the same set is used by the clinical teams at the RSCH and JRH and so was employed in this study to discriminate between *S. epidermidis* infection and *S. epidermidis* specimen contamination (i.e., not infection). *S. epidermidis* was isolated from two or more specimens in the 18/27 (66.7%) cases that were determined to be an ODRI. The 9/27 (33.3%) of cases determined to be not an orthopaedic device-related *S. epidermidis* infection grew a *S. epidermidis* from a single sample only.

One patient case (PDR45) from the John Radcliffe Hospital cultured a *Citrobacter sp.* from 4/5 specimens of a sonicated ankle screw, pus and adjacent tissue and so this was determined to be the cause of ODRI. However, one of the tissue specimens also grew a *S. epidermidis* which was thought to be a contaminant by the clinical team and so was regarded as such for the purposes of this study. For the sake of uniformity throughout this chapter PDR45 is referred to as 'Not Infected' as the *S. epidermidis* isolated was not considered to be the aetiological agent of infection, although the patient clearly had symptoms of an ODRI caused by the *Citrobacter sp.*

The presence of a sinus tract communicating to the site of implant is the other major diagnostic criterion for PJI according to the MSIS. This was only recorded in 3/18 (17%) of patients that had an ODRI. No patient with a sinus tract yielded indistinguishable organisms from more than one peri-operative specimen. This suggests that the presence of a sinus tract, whilst considered indicative of infection, may complicate the interpretation of culture results for surgical samples.

Histological data were not available for intra-operative specimens from the RSCH laboratory as histology was not routinely performed at this site. A diagnosis of non-infection was in complete concordance with a negative histology result (4/4 (100%)) for patient specimens examined at the JRH laboratory. However, 3/17 (18%) diagnoses of infection showed negative histology.

Only 8/15 (53%) of patients who had an ODRI also had purulence at the site of implant recorded, indicating that its absence is not a useful criterion for ruling out infection.

Regarding preoperative diagnostic criteria, patient data were available for serum CRP but not for white cell count. Elevated serum levels of CRP were recorded for patients with an ODRI (16/18 (89%))

as well as for non-infected patients (3/7 (42.9%)). Elevated levels of serum CRP may be associated with several factors, including old age.

Three of nine (33.9%) cases with raised serum CRP levels lacked any other positive diagnostic criteria for infection, resulting in MSIS scores of two. Therefore, they were determined not to be cases of ODRI. Each of these three cases yielded *S. epidermidis* from a single specimen. One of the nine (11.1%) cases with a raised serum CRP level grew *S. epidermidis* from one specimen but also grew a *Citrobacter sp.* from four specimens, so this was determined to be the aetiological agent. One of the nine (11.1%) with a raised serum CRP level grew *S. epidermidis* from one specimen and had purulence at the site of implant, giving an overall score of five according the MSIS criteria. Nevertheless, the clinical team applied the proposition by Parvizi *et al*. (2011) that a single positive culture of a low-virulence organism should not be counted (351). Therefore, the overall MISS score was readjusted to three and the team did not consider this to be a case of ODRI.

		MAJOR CRITERIA		PREOPERATIVE DIAGNOSIS	INTRAOPERATIVE DIAGNOSIS			DECISION	
Patient	**Site**	**>1 positive culture of *indistinguishable organism***	**Sinus tract**	**Elevated Serum CRP (2)**	**Positive histology (3)**	**Positive purulence (3)**	**Single positive culture (2)**	**MSIS Final Score**	**Clinical Decision**
TF12	RSCH	No	No	No	ND	No	Yes	0	Not infected
TF16	RSCH	Yes	No	Yes	ND	Yes			Infected
TF24	RSCH	No	No	No	ND	No	Yes	0	Not infected
TF26	RSCH	Yes	No	Yes	ND	Unknown			Infected
TF32	RSCH	No	No	No	ND	No	Yes*	0	Not infected
TF56	RSCH	No	No	Yes	ND	Yes	Yes	5	Not infected
PDR1	JRH	Yes	No	Yes	Yes	Yes			Infected
PDR6	JRH	Yes	No	Yes	Yes	Yes			Infected
PDR9	JRH	Yes	No	Yes	Yes	Yes			Infected
PDR12	JRH	Yes	No	Unknown	Equivocal	No			Infected
PDR17	JRH	No	No	Yes	No	No	Yes	2	Not infected
PDR19	JRH	Yes	No	Yes	Yes	Unknown			Infected
PDR20	JRH	No	No	Unknown	No	No	Yes	2	Not infected
PDR21	JRH	Yes	No	Yes	Yes	No			Infected
PDR25	JRH	Yes	Yes	Yes	Yes	No			Infected
PDR28	JRH	Yes	No	Yes	Yes	No			Infected
PDR33	JRH	Yes	No	Yes	Yes	No			Infected
PDR36	JRH	Yes	No	Yes	No	No			Infected
PDR38	JRH	Yes	Yes	Yes	No	Unknown			Infected
PDR39	JRH	Yes	No	Yes	ND	Yes			Infected
PDR44	JRH	No	No	Yes	No	No	Yes	2	Not infected
PDR45	JRH	No	No	Unknown	Yes	Yes	Yes	8	Not infected^
PDR49	JRH	No	No	No	No	No	Yes	0	Not infected
PDR93	JRH	Yes	No	Yes	Yes	Yes			Infected
PDR98	JRH	Yes	Yes	Yes	Yes	No			Infected
PDR139	JRH	Yes	No	Yes	Yes	Yes			Infected
PDR144	JRH	Yes	Unknown	Unknown	No	Unknown			Infected

Table 4.3. Table of new scoring-based definition for PJI (Table 1.5) applied to cases of suspected ODRI in which *S. epidermidis* was cultured from one or more specimen. The final decision made by the consultant medical microbiologist is included in the final column. MSIS scores for pre-operative and peri-operative minor criteria are shown in brackets. **S. epidermidis* was isolated from the enrichment broth of one specimen. ^*Citrobacter sp.* was isolated from 4 specimens and so was determined to be the causative organism, *S. epidermidis* was isolated from 1 specimen so determined to be a contaminant. CRP: C-Reactive Protein. WBC: White Blood Count. LE: Leukocyte Esterase. ND: Not Done.

4.4.2 Genomic and Phenotypic Analyses

To investigate the population structure of *S. epidermidis* associated with cases of ODRI and with specimen culture contamination five colonies from positive cultures were analysed. Five individual colonies from primary culture plates were sub-cultured onto a CBA plate (Oxoid Ltd, Basingstoke, UK) and incubated overnight aerobically at 37°C to give enough cultured organism for further investigations. Henceforth each sub-cultured colony will be referred to as an isolate. In total 325 isolates from 65 specimen cultures were analysed. Table 4.4. shows the quality data of the reads generated by HiSeq sequencing for all the isolates. In addition, 10/325 (3.1%) isolates were randomly selected to be sequenced in duplicate to accord with MMM and WTCHG protocols, to detect false-positive variant calling. Table 4.5. shows that the ten duplicates were all 0 SNVs from the originals, consistent with the previously published estimate of approximately 2.5 x 10^{-9} errors per nucleotide using the MMM pipeline (285). Duplicates were removed from downstream analyses.

Parameter	Range	Median (IQR)
Reads (n)	1.44 – 6.51 x 10^6	3.83 x 10^6 (3.39 – 4.48 x 10^6)
Mapped (n)	1.31 – 6.35 x 10^6	3.47 x 10^6 (3.03 – 4.10 x 10^6)
Mapped (%)	79.0 – 97.62	91.51 – (88.82 – 93.77)

Table 4.4. Quality control data for 325 clinical isolate genome sequences

	PDR19bm5.2	PDR21bm5.2	PDR28cm3.2	PDR38bm1.2	PDR93am4.2	PDR98am1.2	PDR139bm3.2	PDR144am2.2	TF29m2.2	TF51m1.2
PDR19bm5	0	6891	7129	6214	8732	5876	5045	54602	54334	5392
PDR21bm5	6891	0	7838	6490	9343	6014	7488	54259	54189	6240
PDR28cm3	7129	7838	0	7727	2547	7590	6799	54738	54781	8436
PDR38bm1	6214	6490	7727	0	8277	4925	7302	54167	53865	3627
PDR93am4	8732	9343	2547	8277	0	8201	6841	53486	53417	8588
PDR98am1	5876	6014	7590	4925	8201	0	6354	54203	54264	4229
PDR139bm3	5045	7488	6799	7302	6841	6354	0	53653	53541	6633
PDR144am2	54602	54259	54738	54167	53486	54203	53653	0	9694	53297
TF29m2	54334	54189	54781	53865	53417	54264	53541	9694	0	53305
TF51m1	5392	6240	8436	3627	8588	4229	6633	53297	53305	0

Table 4.5. Pairwise distance matrix showing the single-nucleotide variation between original and duplicate samples that were whole-genome sequenced in duplicate to ensure accuracy of data. 10/325 samples were sequenced twice and a pairwise distance of SNV differences in the genome were measured. There were 0 SNVs between original (rows) and duplicate (columns) isolate genomes when mapped to reference genome ATCC 12228 therefore there was limited chance of SNVs being falsely called.

4.4.3 Antimicrobial Susceptibility Testing and Multi-Locus Sequence Typing

Each of the 325 clinical isolates (Table 4.6 and Table 4.7), as well as five colonies of the reference strain NCTC 11047 (Table 4.8) underwent the AST procedure explained in Chapter 2: Methods. The 13 antibiotics tested were the same as those tested by the microbiology laboratory at the Royal Sussex County Hospital. The set of antibiotics tested aims to provide the clinicians with a range of treatment options for deep and superficial staphylococcal infection. Multi-locus sequence types were determined *in silico* from whole-genome sequence data (see Chapter 2: Methods).

4.4.3.1 Antimicrobial Susceptibility Profiles

Amongst the 325 isolates from 65 specimen cultures a total of 25 different antibiograms were observed. Three of the antibiograms were observed in two different sequence types and one antibiogram was observed in three different STs showing that antibiotic resistance profiling is not always a reliable indicator for determining the relatedness of isolates of *S. epidermidis*. Further analyses of antimicrobial susceptibility are documented in Chapter 5: Detecting genomic determinants of antimicrobial resistance in *Staphylococcus epidermidis*.

4.4.3.2 *In silico* multi-locus sequence types

A total of 18 different STs were detected amongst the clinical isolates of *S. epidermidis*. ST2 was the most frequently detected, from 8/27 (29.6%) of cases. This finding concurs with previous studies that found ST2 to be a prevalent hospital-adapted strain either and often the most common or amongst the most common ST to be detected from clinical specimens (32,120,133,238,301). Moreover, ST2 was detected both as an infecting and contaminating isolate which concurs with previous studies that show that nosocomial strains may be isolated equally as a cause of infection or a result of specimen contamination by hospital-associated commensal bacteria (109,128).

Isolates of the same sequence type showed marked diversity according to the antibiogram results. There were eight different antibiotic phenotypes amongst the ST2 isolates, including one patient (PDR93) who had two different strains according to the antibiograms. Similarly, ST57 and ST88 were both isolated from more than one patient that showed different antibiograms. This demonstrates that neither multi-locus sequence type nor antibiograms give a reliable information about the diversity or relatedness of clinical isolates of *S. epidermidis*.

Table 4.9 shows that of the 18 detected STs six had novel MLST profiles, of which half were a result of a novel gene locus or lack of locus. This suggests that there are still further nosocomial strains of *S. epidermidis* that remain to be detected and classified.

4.4.3.3 Relatedness of Isolates from the Same Specimen Set According to AST and MLST

For 14/18 cases of ODRI all colonies in a specimen set showed identical antibiograms and MLST profiles meaning they were indistinguishable by conventional methods. Four of the 18 ODRI cases cultured different colonies either by antibiogram, MLST or both.

One specimen from patient TF26 grew colonies of ST588 which had the same antibiogram and two specimens grew colonies of a previously unrecorded sequence type which was given the designation ST772. All colonies of ST772 had the same antibiogram as each other.

The two specimens analysed from patient PDR25 grew colonies of *S. epidermidis* which were indistinguishable by antibiogram however two different sequence types were detected from WGS data. The first was ST2, one of the most common nosocomial sequence types (32,120,133,238,301). A second, novel sequence type was also detected that was lacking the *arcC* locus. The *arcC* gene is part of the *arcABDCR* operon and encodes a carbamate kinase which is part of the main arginine catabolic pathway in *S. aureus* (352). The *arcC* gene is thought to be a housekeeping gene, essential for bacterial metabolism (353) and was proposed originally by Wisplinghoff *et al.* (2003) then by Thomas *et al.* (2007) to form part of their MLST schemes for *S. epidermidis* (246,353). The novel sequence type was submitted to the pubMLST.org database and was given the designation ST781. The genomes of the ST781 isolates were interrogated with BLASTn (see Chapter 2: Methods) for the presence of the *arcC* gene but it was not detected meaning it was not just the locus used for MLST that was not present but the entire gene.

Patient PDR93 had two specimens selected for analysis, both of which grew ST2 *S. epidermidis* however there were two different antibiograms observed suggesting different strains of the same sequence type.

Patient	Specimen	Specimen	Cefox	Pen	Eryth	Gent	Vanc	Teic	Fus	Rif	Tet	Cipro	Clind	Lzd	Mup	MLST
TF12	TF12a	Screw from DHS	I	R	S	S	S	S	R	S	S	S	S	S	S	208
TF16	TF16a	Deep tissue	R	R	R	S	S	S	R	S	S	S	S	S	S	640
	TF16b	Deep tissue	R	R	R	S	S	S	R	S	S	S	S	S	S	640
	TF16c	Deep tissue	R	R	R	S	S	S	R	S	S	S	S	S	S	640
TF24	TF24a	Acetabulum	R	R	R	R	S	S	S	S	S	R	R	S	S	2
TF26	TF26a	Tissue	R	R	R	S	S	S	R	R	S	R	R	S	R	772
	TF26b	Tissue	S	R	R	S	S	S	S	S	S	S	S	S	R	558
	TF26c	Tissue	R	R	R	S	S	S	R	R	S	R	R	S	R	772
TF32	TF32a	Tissue	S	R	S	S	S	S	S	S	S	S	S	S	S	88
TF56	TF56a	Tissue	S	R	S	S	S	S	S	S	S	S	S	S	S	782
PDR1	PDR1a	Fluid	R	R	S	R	S	S	S	S	S	S	S	S	S	89
	PDR1b	Capsule	R	R	S	R	S	S	S	S	S	S	S	S	S	89
	PDR1c	Fluid	R	R	S	R	S	S	S	S	S	S	S	S	S	89
	PDR1d	Femoral membrane	R	R	S	R	S	S	S	S	S	S	S	S	S	89
PDR6	PDR6a	Synovium	I	S	S	S	S	S	R	S	S	S	S	S	S	88
	PDR6b	Patella membrane	I	S	S	S	S	S	R	S	S	S	S	S	S	88
	PDR6c	Femoral membrane	I	S	S	S	S	S	R	S	S	S	S	S	S	88
	PDR6d	Tibial membrane	I	S	S	S	S	S	R	S	S	S	S	S	S	88
	PDR6e	Joint fluid	I	S	S	S	S	S	R	S	S	S	S	S	S	88
PDR9	PDR9a	Joint fluid	I	R	R	S	S	S	R	S	S	S	S	S	S	35
	PDR9b	Femoral membrane	I	R	R	S	S	S	R	S	S	S	S	S	S	35
	PDR9c	Anterior capsule	I	R	R	S	S	S	R	S	S	S	S	S	S	35
PDR12	PDR12a	Capsule	I	R	R	S	S	S	R	S	S	R	S	S	S	2
	PDR12b	Neck Capsule	I	R	R	S	S	S	R	S	S	R	S	S	S	2
	PDR12c	Acetabular scar tissue	I	R	R	S	S	S	R	S	S	R	S	S	S	2
	PDR12d	Acetabular membrane	I	R	R	S	S	S	R	S	S	R	S	S	S	2
	PDR12e	Acetabular membrane	I	R	R	S	S	S	R	S	S	R	S	S	S	2
PDR17	PDR17a	Deep inferior part	R	R	R	R	S	S	R	S	S	S	R	S	S	2
PDR19	PDR19a	Joint fluid	S	S	S	S	S	S	R	S	S	S	S	S	S	958
	PDR19b	Capsule	S	S	S	S	S	S	R	S	S	S	S	S	S	958
	PDR19c	Synovium	S	S	S	S	S	S	R	S	S	S	S	S	S	958

	PDR19d	Femoral membrane	S	S	S	S	S	S	R	S	S	S	S	S	S	958
PDR20	PDR20a	Acetabular membrane	S	S	R	S	S	S	S	S	S	S	S	S	S	764
PDR21	PDR21a	Superficial tissue	R	R	S	S	S	S	S	S	S	S	S	S	S	88
	PDR21b	Superfascia fluid	R	R	S	S	S	S	S	S	S	S	S	S	S	88
	PDR21c	Superficial fluid	R	R	S	S	S	S	S	S	S	S	S	S	S	88
	PDR21d	Deep tissue	R	R	S	S	S	S	S	S	S	S	S	S	S	88
	PDR21e	Anterior capsule	R	R	S	S	S	S	S	S	S	S	S	S	S	88
PDR28	PDR28a	Proximal femur	R	R	R	R	S	S	R	S	S	R	S	S	R	2
	PDR28b	Tibial membrane	R	R	R	R	S	S	R	S	S	R	S	S	R	2
	PDR28c	Distal tibia tissue	R	R	R	R	S	S	R	S	S	R	S	S	R	2
PDR33	PDR33a	Hip membrane	R	R	R	R	S	S	S	R	R	S	R	S	S	2
	PDR33b	Acetabular tissue	R	R	R	R	S	S	S	R	R	S	R	S	S	2
PDR36	PDR36a	Joint fluid	R	R	R	R	S	S	S	S	S	R	R	S	R	2
	PDR36b	Femoral membrane	R	R	R	R	S	S	S	S	S	R	R	S	R	2
PDR39	PDR39a	Knee fluid	R	R	R	R	S	S	R	S	S	R	R	S	S	87
	PDR39b	Tibial tunnel tissue	R	R	R	R	S	S	R	S	S	R	R	S	S	87
PDR44	PDR44	Hip synovium	S	R	R	S	S	S	R	S	S	S	S	S	S	208
PDR45	PDR45	Tissue	S	R	R	S	S	S	R	S	S	S	S	S	S	7
PDR49	PDR49a	Sequestrum	S	R	S	S	S	S	S	S	S	S	S	S	S	89
PDR98	PDR98a	Hip sinus	S	R	R	R	S	S	S	R	R	R	S	S	S	57
	PDR98b	Deep hip sinus	S	R	R	R	S	S	S	R	R	R	S	S	S	57
PDR139	PDR139a	Capsule	R	R	R	R	S	S	R	S	R	R	R	S	S	5
	PDR139b	Femoral neck tissue	R	R	R	R	S	S	R	S	R	R	R	S	S	5
	PDR139c	Superficial hip fascia	R	R	R	R	S	S	R	S	R	R	R	S	S	5
PDR144	PDR144a	Tissue	S	R	R	S	S	S	S	S	S	S	S	S	S	924
	PDR144b	Tissue	S	R	R	S	S	S	S	S	S	S	S	S	S	924

Table 4.6. Antimicrobial susceptibility testing and MLST results of 5 colonies tested from *S. epidermidis* positive cultures. Five colonies were tested individually from each culture and showed identical results. S: Susceptible, R: Resistant, I: Intermediate, Cefox: Cefoxitin, Pen: Penicillin, Eryth: Erythromycin, Gent: Gentamicin, Vanc: Vancomycin, Teic: Teicoplanin, Fus: Fusidic Acid, Rif: Rifampicin, Tet: Tetracycline, Cipro: Ciprofloxacin, Clind: Clindamycin, Lzd: Linezolid, Mup: Mupirocin.

Patient	Colony	Specimen	Cefox	Pen	Eryth	Gent	Vanc	Teic	Fus	Rif	Tet	Cipro	Clind	Lzd	Mup	MLST
PDR25	PDR25a1	Acetabular membrane	R	R	S	R	S	S	R	R	S	R	S	S	S	781
	PDR25a2		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR25a3		R	R	S	R	S	S	R	R	S	R	S	S	S	781
	PDR25a4		R	R	S	R	S	S	R	R	S	R	S	S	S	781
	PDR25a5		R	R	S	R	S	S	R	R	S	R	S	S	S	781
	PDR25b1	Fascia layer	R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR25b2		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR25b3		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR25b4		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR25b5		R	R	S	R	S	S	R	R	S	R	S	S	S	2
PDR38	PDR38b1	Tissue 2	S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38b2		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38b3		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38b4		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38b5		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38c1	Tissue 3	S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38c2		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38c3		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38c4		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38c5		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38d1	Tissue 4	S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38d2		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38d3		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38d4		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38d5		S	S	R	S	S	S	S	S	S	S	S	S	S	57
	PDR38e1	Tissue 5	S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38e2		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38e3		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38e4		S	S	R	S	S	S	R	S	S	S	S	S	S	130
	PDR38e5		S	S	R	S	S	S	R	S	S	S	S	S	S	130

PDR93	PDR93a1	Capsule	R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93a2		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93a3		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93a4		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93a5		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93b1	Femoral membrane	R	R	R	R	S	S	R	R	S	R	R	S	R	2
	PDR93b2		R	R	R	R	S	S	R	R	S	R	R	S	R	2
	PDR93b3		R	R	S	R	S	S	R	R	S	R	S	S	S	2
	PDR93b4		R	R	R	R	S	S	R	R	S	R	R	S	R	2
	PDR93b5		R	R	R	R	S	S	R	R	S	R	R	S	R	2

Table 4.7. Antimicrobial susceptibility testing and MLST results of individual colonies tested from *S. epidermidis* positive cultures. Five colonies were tested individually from each culture and showed differing results. S: Susceptible, R: Resistant, I: Intermediate, Cefox: Cefoxitin, Pen: Penicillin, Eryth: Erythromycin, Gent: Gentamicin, Vanc: Vancomycin, Teic: Teicoplanin, Fus: Fusidic Acid, Rif: Rifampicin, Tet: Tetracycline, Cipro: Ciprofloxacin, Clind: Clindamycin, Lzd: Linezolid, Mup: Mupirocin.

Strain	Cefox	Pen	Eryth	Gent	Vanc	Teic	Fus	Rif	Tet	Cipro	Clind	Lzd	Mup	MLST
NCTC 11047	S	R	S	S	S	S	R	S	R	S	S	S	S	5

Table 4.8. Antimicrobial susceptibility testing and MLST results of 5 colonies tested of the NCTC 11047 *S. epidermidis* type strain. Five colonies were tested individually and showed identical results. S: Susceptible, R: Resistant, Cefox: Cefoxitin, Pen: Penicillin, Eryth: Erythromycin, Gent: Gentamicin, Vanc: Vancomycin, Teic: Teicoplanin, Fus: Fusidic Acid, Rif: Rifampicin, Tet: Tetracycline, Cipro: Ciprofloxacin, Clind: Clindamycin, Lzd: Linezolid, Mup: Mupirocin.

MLST	Allelic Profile							Specimen
	arcC	*aroE*	*gtr*	*mutS*	*pyrR*	*tpiA*	*yqiL*	
764	28	41	15	5	38	34	4	Contaminant
772	1	1	1	2	2	1	**62**	ORDI
781	**0**	1	2	2	4	1	1	ORDI
782	16	17	2	1	2	1	1	ODRI
924	12	29	5	5	3	33	4	ODRI
958	1	**76**	1	2	5	1	1	ODRI

Table 4.9. Novel MLST allele profiles submitted and accepted into the pubMLST BIGSdb *Staphylococcus epidermidis* database during this study. Novel alleles are highlighted in bold.

4.4.1 Whole-genome Sequencing Pairwise Distance Analysis of Isolates from the Same Specimen and Specimen Set

Young *et al.* (2017) observed that, on average, mean pairwise distance was lower amongst genomes of infecting isolates of *S. aureus* ($\pi = 0.6$) than nasal carriage isolates ($\pi = 2.8$) (318). As illustrated in Figure 4.2., I hypothesise that this observation could be used to develop a method that could be used to discriminate between infecting and contaminating isolates of *S. epidermidis*. It is likely that the cases of ODRI that result from inoculation of the orthopaedic device at the time implantation are a bottleneck event – where a small number of highly-related bacterial cells are involved. Subsequently these cells undergo division and diversify, accumulating SNVs over time. A similar process is likely for specimen or culture contamination except that the shorter timeframe between initial inoculation and culture means that whilst several SNVs will have accumulated in infecting populations a much smaller number (or none) will have accumulated in contaminating isolates.

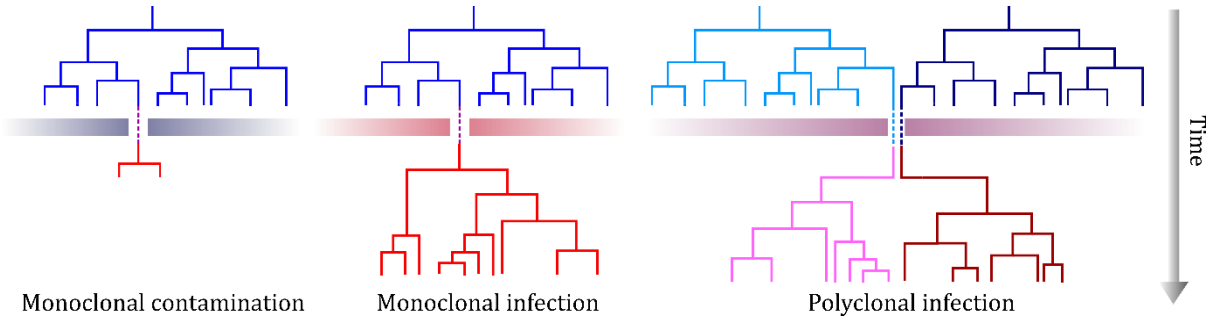


Figure 4.2. Working hypothesis of the difference in diversity between monoclonal contamination, monoclonal infection and polyclonal infection. The thick horizontal bars represent the inoculation event either during implantation (for infection) or specimen explantation/processing (for contamination). Adapted from Young *et al.* (2017) (318).

To determine if this hypothesis could be applied to discriminate between infecting and contaminating *S. epidermidis* the WGS data of isolates from individual specimens and from specimen sets underwent pairwise analysis. Firstly, raw reads were mapped to the Sanger-sequenced reference genome of *S. epidermidis* ATCC 12228 using the Stampy algorithm to generate whole-genome sequences formed of contigs (283). The ClonalFrameML tool was then used to determine the number of SNVs between each genome sequence whilst detecting recombination (311). There was no recombination detected therefore diversity was a result of SNV accumulation alone. Pairwise

distance matrices were constructed from the number of SNVs separating each genome. The mean pairwise distance was calculated as the sum of all distances in - a) a single specimen or b) all specimens in a set - divided by the number of non-self distances. For example, Figure 4.3 shows that the sum of all pairwise distances of genomes from specimen PDR9a was 32 and there were ten non-self pairwise distances giving a mean pairwise distance of 3.2.

		PDR9a					PDR9b					PDR9c				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
PDR9a	1															
	2	1														
	3	2	1													
	4	6	1	6												
	5	6	1	6	2											
PDR9b	1	23	19	23	21	21										
	2	23	19	23	21	21	0									
	3	23	19	23	21	21	0	0								
	4	23	19	23	21	21	0	0	0							
	5	23	19	23	21	21	0	0	0	0						
PDR9c	1	5	0	5	1	1	20	20	20	20	20					
	2	6	1	6	2	2	21	21	21	21	21	1				
	3	5	0	5	1	1	20	20	20	20	20	0	1			
	4	6	1	6	2	2	21	21	21	21	21	1	0	1		
	5	7	2	7	3	3	22	22	22	22	22	2	1	2	1	

Figure 4.3. Example of the pairwise distance matrix for genomes from case PDR9. The number of SNVs separating the genomes of 5 colonies from each specimen: PDR9a, PDR9b and PDR9c were input. Yellow shaded squares were used to calculate the mean pairwise distance between colonies from the same specimen. Purple shaded squares were used to calculate the mean pairwise distance between colonies from different specimens in the set.

There was a marked difference between the mean pairwise distance of genomes from the infection-associated culture sets and the contamination-associated culture sets. The highest mean pairwise distance observed in the contaminant-associated genomes was 0.4 SNVs in specimen TF12 (Figure 3.5). The highest mean pairwise distance of genomes of the same sequence type and antibiogram observed in the infection-associated culture sets was 17.64 SNVs in culture-set PDR33. Culture-sets that had more than one different sequence type had mean pairwise distances in the thousands showing that polyclonal infection can contribute to markedly high diversity. Culture-set PDR93 grew isolates of the same sequence type however two different phenotypes were observed according to the antibiogram. All ten isolates were ST2 however four were resistant to erythromycin, clindamycin

and mupirocin and six were susceptible. There was a mean pairwise distance of 34 SNVs across the ten isolates demonstrating that they were genomically closely related. Interrogation of the genomes showed that the four colonies resistant to erythromycin, clindamycin and mupirocin harboured the genes *ermC* and *mupA*, both known resistance determinants whilst the six susceptible colonies lacked these genes. The reference genome ATCC 12228 harbours neither *mupA* nor *ermC* and therefore were not present in the mapped genomes mapped meaning that these differences were not observed. Conversely, culture-set PDR25 had isolates of two different sequence types (ST2 and ST781) but all had the same antimicrobial phenotype. The mean pairwise distance was only 0.15 SNVs. As mentioned previously the difference in sequence type was accounted for by the lack of the *arcC* gene in the ST781 isolates. Again, the low pairwise distance despite the lack of an entire gene is a consequence of the fact that ClonalFrameML (311)allows for gene deletion and shows that the two sets of isolates are highly related genomically. Whilst the loss of entire genes results in many differences at the nucleotide level of the genome, they can represent a single evolutionary event and were therefore treated as such in the analysis.

The Mann Whitney U Test (Wilcoxon Rank-Sum Test) was calculated using R (354) to determine the statistical difference in diversity between the infection-associated and contamination-associated culture-sets. The p-value was $p = 0.000115$. However, this p-value is only an approximation due to the high proportion of tied observations (there were 35 zero-SNV observations). Since all the contamination-associated culture-sets had only a single culture whilst all the infection-associated culture-sets had more than one culture a fairer way of determining the difference between the two would be to analyse the genomes from individual cultures.

The highest mean pairwise distance of genomes of the same sequence type and phenotype observed in the infection-associated cultures was 12.4 SNVs in specimen PDR33a. Specimen PDR38c had a mean pairwise distance of 4285 SNVs however this was due to there being two different sequence types isolated from the same specimen. The Wilcoxon Rank-Sum Test (Mann-Whitney U Test) was used to determine the statistical difference between the contamination-associated and infection-associated specimens. The p-value between the two sets of mean pairwise distances was $p = 0.0254$ meaning that the null hypothesis can be rejected.

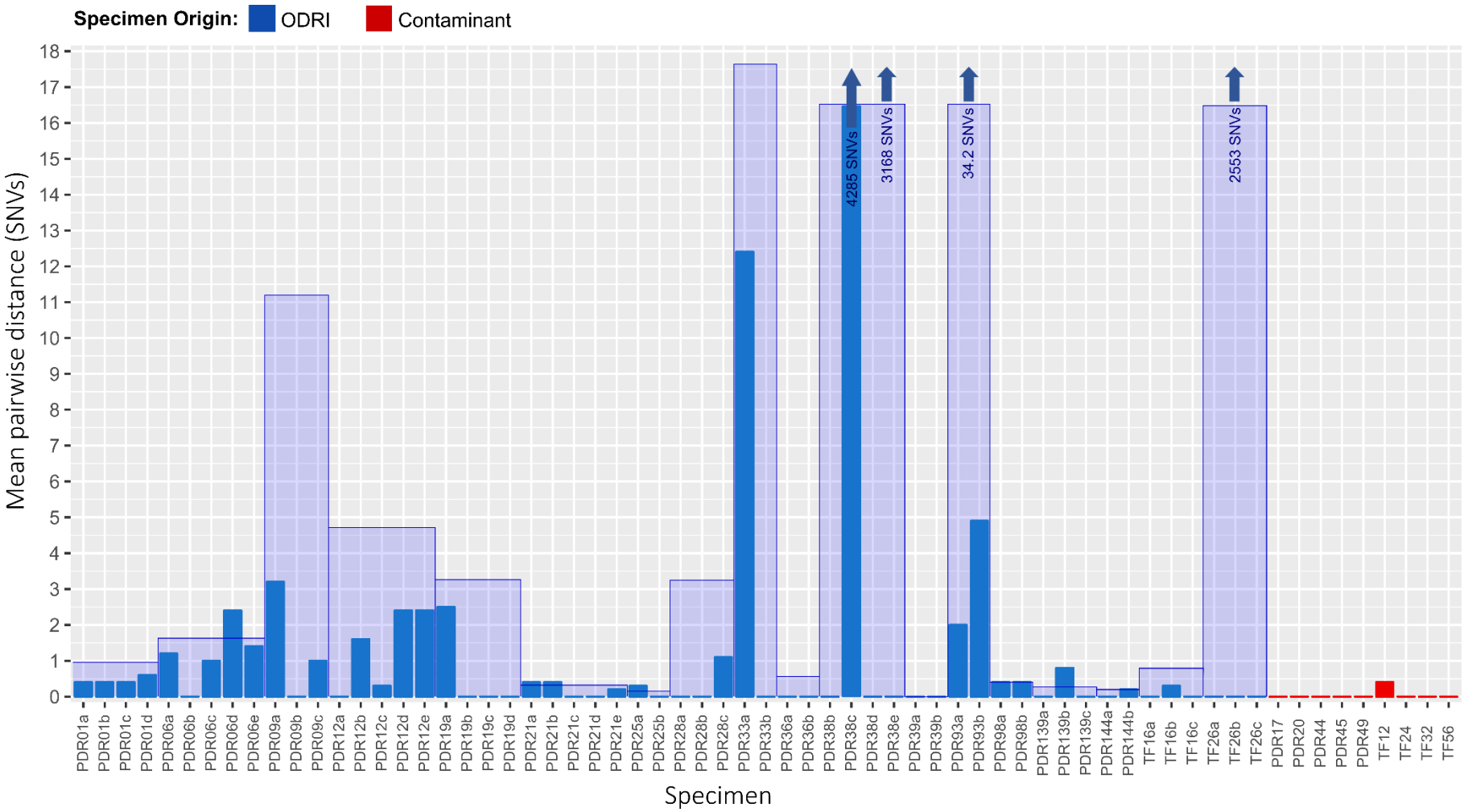


Figure 4.4. Bar chart of the mean pairwise distance of genomes of *S. epidermidis* from individual cultures (solid bars) and culture sets (translucent bars). True orthopaedic device-related infection-associated are blue and contamination-associated cultures are red. Mean pairwise distances greater than 18 SNVs are indicated with an arrow and are associated with more than one sequence type or antimicrobial phenotype being isolated from the same culture or culture-set.

4.4.2 Haplotrees to characterise the diversity of *S. epidermidis* in infection and contamination

To visualise the diversity of *S. epidermidis* ODRI and specimen contamination haplotrees were constructed from genomic data. This was done using a method implemented by Professor Daniel Wilson and used by Bernadette Young *et al.* (2017) to characterise mutations associated with the change from commensal nasally-colonising to disease-causing *S. aureus* (318). For this study, the method was applied to observe the genotypes detected amongst the colonies of *S. epidermidis* analysed for each case of ODRI or specimen contamination. The genome of each colony was compared with the reference genome of *S. epidermidis* ATCC 12228 to identify the location of any SNVs that existed between the reference genome and the subject genomes. These positions were then used to construct haplotrees which showed the observed genotypes as a white/grey nodes or unobserved hypothetical genotypes as smaller black nodes.

Figure 4.5a shows the haplotrees constructed from three cases of ODRI in which five specimens growing *S. epidermidis* were analysed. Case *PDR6* was the first stage of a TKR revision. The haplotree from *PDR6* shows a central genotype that was detected in four of the specimens from which other genotypes radiate at a maximum of two SNVs. This central genotype may be a dominant clone more populace than the others at the site of infection. Therefore, the central genotype could potentially represent the initially infecting clone, or a prevailing clone better adapted to the ecological niche and/or insult by the host immune system or antimicrobial agents. The lateral nodes may represent genotypes that are the result of natural diversification which occurs over time. The nine radiating edges show that a single genotype has given rise to nine different sub-lineages. Case *PDR12* is from a THR and has a haplotree with a central genotype detected multiple times in more than one specimen. However, the radiating edges are longer with up to eight SNVs representing a more divergent population structure. *PDR21* was a case of the DAIR of an infected THR. The haplotree is much more clonal with only two genotypes detected which differed by a single SNV. This markedly lower diversity may be the result of this case being an early infection in which the population has had less time to acquire SNVs or a selective sweep reducing the diversity of the population. The DAIR procedure is more effectively applied during early infection (270) and so this may also be an indicator of a more recent inoculation event.

Figure 4.5b is comprised of haplotrees constructed from cases of ODRI in which *S. epidermidis* from four specimens were analysed. Case *PDR1* was an infected hemiarthoplasty leading to a single stage THR. The haplotree for *PDR1* shows a central genotype isolated from 3/4 specimens. From the central genotype radiate three genotypes differing by a single SNV from the same specimens and two genotypes from the remaining specimen one of which differs from the central genotype by a

single SNV. This again suggests a closely related clonal population expanding from an initially colonising or prevailing clone from which the sublineages diverge. Case *PDR19* had six specimens growing *S. epidermidis*, four of which were analysed for this study. The associated haplotree shows an inferred unobserved genotype linking three observed genotypes. This suggests there was a central genotype that was either undetected in this study or supplanted by more successful variants. A single genotype was isolated from the synovial fluid and the synovial membrane and could be due to the immediate proximity of the two in the joint. Case *PDR38* involved infection of acetabular metalwork. There were two different sequence types (*ST57* and *ST130*) isolated from the specimens which differed by an average of 3168 SNVs (see Figure 4.4) and had different antimicrobial susceptibility phenotypes (see Table 4.7). ST57 was isolated from three specimens and differed by eight SNVs. Colonies of ST130 were isolated from two specimens and had no SNVs separating them. Both sequence types were isolated from multiple specimens strongly suggesting polyclonal infection. It is possible that both STs were present in all specimens but were not detected however this is not known.

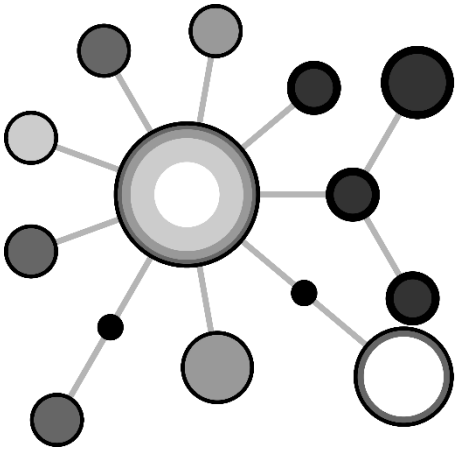
Figure 4.5c is comprised of haplotrees constructed from cases in which three specimens growing *S. epidermidis* were analysed. The haplotree for case *PDR9* shows a slightly radial structure with 7 genotypes being detected in two specimens and having a maximum divergence of 22 SNVs. A more distantly related genotype was detected in a single specimen. *PDR28* has a slightly radial structure with a single specimen growing two genotypes and the other two growing a single genotype each. *PDR139* shows a more homogeneous population structure with only two SNVs separating all 15 genomes. This may be due to the infection being in its early stage as suggested by the associated DAIR treatment. Case *TF16* is more homogeneous still with only two SNVs separating all genomes. The two genotypes shown in the haplotree for case *TF26* is separated by 7598 SNVs and is the result of two different sequence types (ST558 isolated from one specimen and ST772 isolated from two specimens) with different antimicrobial susceptibility phenotypes. Case *TF26* had 5/5 specimens growing *S. epidermidis* however for this study only three were analysed therefore it is not known if ST558 was present in the other specimens suggesting it was part of a polyclonal infection or only present in one specimen indicating it was a contaminant.

Figure 4.5d shows the haplotrees of eight cases of *S. epidermidis* ODRI from which two specimens per case were analysed. Of the eight cases, five involved the hip, two involved the knee and one involved the calcaneus. *PDR25* shows a haplotree consisting of a single genotype isolated from one specimen and two genotypes separated from each other by one SNV isolated from another specimen. *PDR33* shows a single genotype isolated from one single specimen and two genotypes 29 SNVs distant from each other and 25/26 SNVs distant from the first genotype. Haplotree *PDR36* is

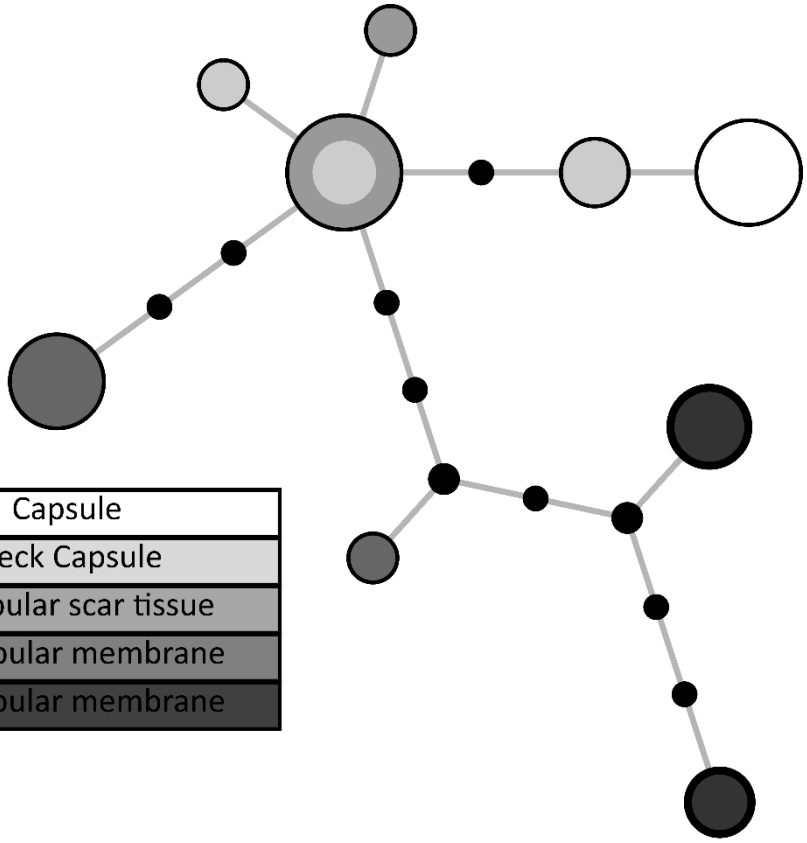
comprised of one genotype isolated from each specimen and only differing by a single SNV. Haplotree *PDR39* is uninodal representing ten completely homogeneous genomes being isolated from both specimens. *PDR9*3 is a more complex haplotree comprised of eight observed genotypes (three from one specimen and five from the other specimen). Genotypes isolated from the two different specimens in the set are separated by a minimum of 50 SNVs. Whilst all the genomes from case *PDR93* are of the same sequence type there are two different antimicrobial susceptibility phenotypes. Only six SNVs separate the two genotypes with different antimicrobial susceptibility phenotypes. Table 4.7 shows that 6/10 colonies from case *PDR93* were resistant to cefoxitin, penicillin, gentamicin, fusidic acid, rifampicin and ciprofloxacin and 4/10 were also resistant to erythromycin, clindamycin and mupirocin. Analysis of the *de novo* assembled genomes found that the four more resistant genomes harboured *ermC* and *mupA* whilst the six less resistant colonies did not. Subsequent interrogation with PlasmidFinder (355) found the *mupA* and *ermC* genes were respectively located on plasmid repUS19 and a plasmid with 100% identity to pNE131, both which were not present in the genomes of the six erythromycin, clindamycin and mupirocin susceptible colonies. This demonstrates that plasmid loss or gain, whilst is a significant genomic event could result in closely related isolates being mistaken for distant relatives. The haplotree for case *PDR98* shows three genotypes one of which was isolated from the two specimens. The case for *PDR144* involved the peri-prosthetic tissue surrounding a screw employed to fix a calcaneal fracture. In total *S. epidermidis* was cultured from 4/5 of the specimens explanted for this case, however only two underwent genomic analyses. Of the ten colonies isolated, nine had the same genotype and one colony differed by a single SNV. The specimens cultured from case *TF51* were not explanted during surgery but were taken prior to surgery. One of the specimens from *TF51* was a processed as a superficial wound swab and the *S. epidermidis* that was cultured was reported as skin flora, the other specimen was a synovial fluid which was reported as a significant isolate with associated AST results for administration of antibiotic therapy. Seven of the colonies of *S. epidermidis* cultured from the two specimens were genomically identical and the other three were very closely related indicating that these were the same strain.

Figure 4.5e shows nine haplotrees constructed from the genomes of *S. epidermidis* determined by the consultant medical microbiologist's application of the MSIS criteria to be contaminants of specimen cultures from cases of suspected ODRI. Since the MSIS criteria consider more than one specimen growing an indistinguishable organism to be a major indicator of infection (76) each case had only a single specimen that was culture-positive therefore each haplotree is constructed from five genomes only. 8/9 of the haplotrees were uninodal meaning no SNVs were detected between

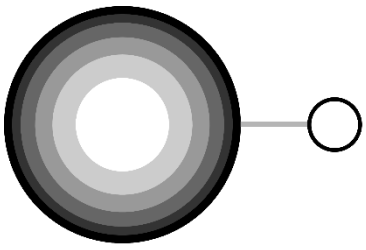
the five genomes. The remaining haplotree shows a single genotype isolated four times and another genotype one SNV different isolated once.



PDR6a	Synovium
PDR6b	Patella membrane
PDR6c	Femoral membrane
PDR6d	Tibial membrane
PDR6e	Joint fluid

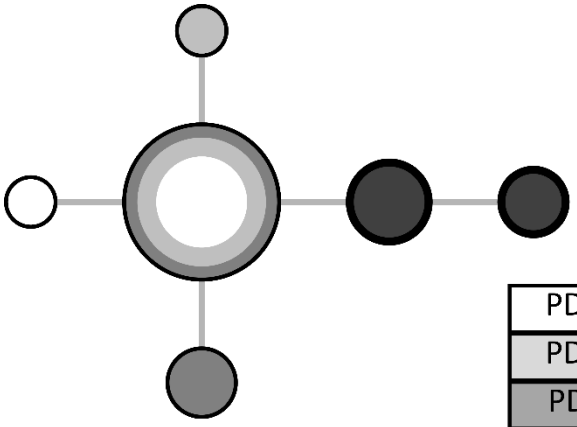


PDR12a	Capsule
PDR12b	Neck Capsule
PDR12c	Acetabular scar tissue
PDR12d	Acetabular membrane
PDR12e	Acetabular membrane

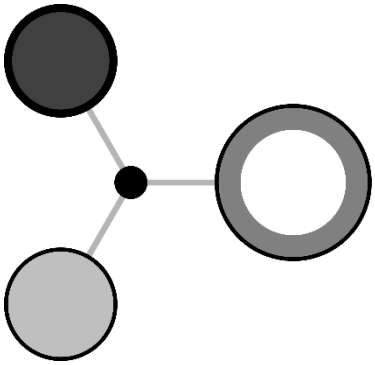


PDR21a	Superficial tissue
PDR21b	Superfascia fluid
PDR21c	Superficial fluid
PDR21d	Deep tissue
PDR21e	Anterior capsule

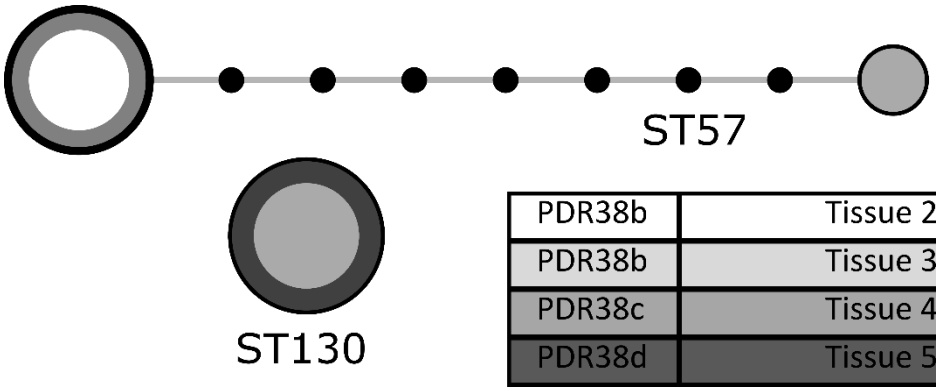
Figure *4.5*a. Haplotrees drawn from the detected genotypes of *S. epidermidis* colonies cultured from ODRI. The size of each node is proportional to the number of times a genotype was detected. Black nodes represent unobserved intermediate hypothetical genotypes. Grey edges represent the mutations that separate genotypes White/grey shading indicates specimen.



PDR1a	Fluid
PDR1b	Capsule
PDR1c	Fluid
PDR1d	Femoral membrane

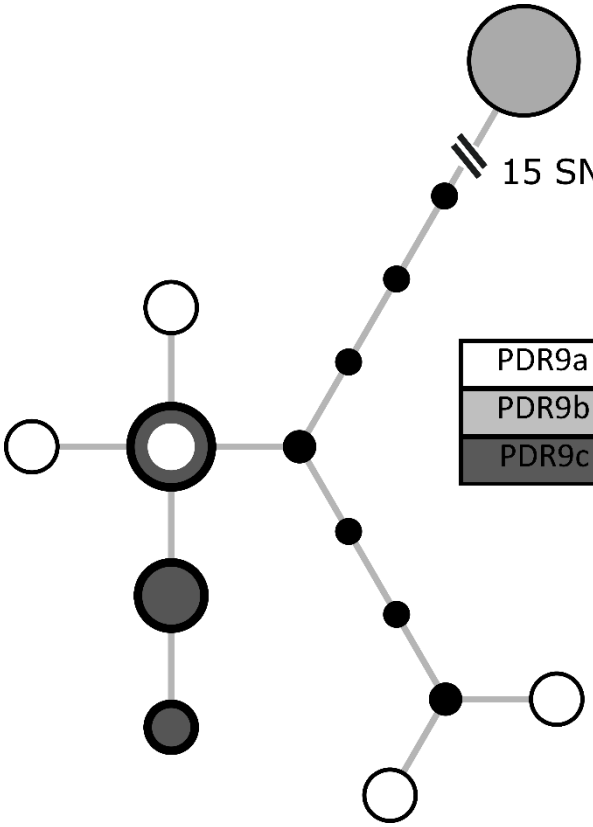


PDR19a	Joint fluid
PDR19b	Capsule
PDR19c	Synovium
PDR19d	Femoral membrane

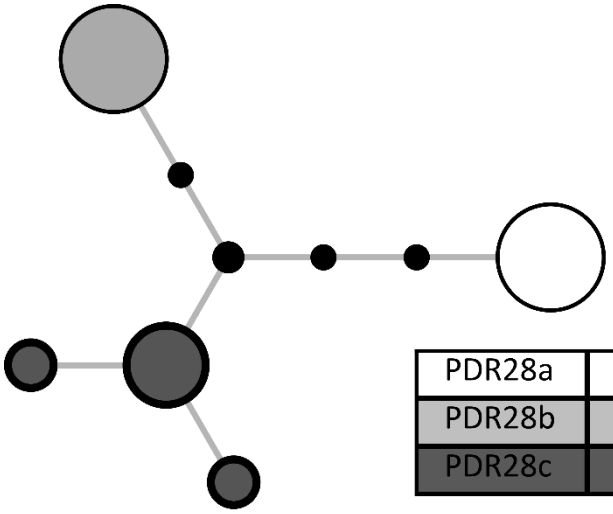


PDR38b	Tissue 2
PDR38b	Tissue 3
PDR38c	Tissue 4
PDR38d	Tissue 5

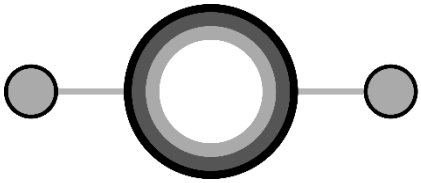
Figure *4.5*b. Haplotrees drawn from the detected genotypes of *S. epidermidis* colonies cultured from ODRI. The size of each node is proportional to the number of times a genotype was detected. Black nodes represent unobserved intermediate hypothetical genotypes. Grey edges represent the mutations that separate genotypes White/grey shading indicates specimen.



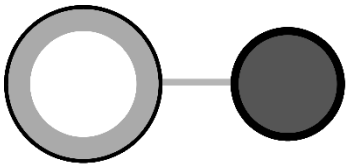
PDR9a	Joint fluid
PDR9b	Femoral membrane
PDR9c	Anterior capsule



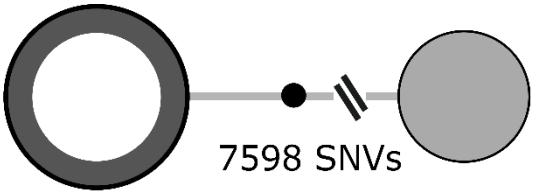
PDR28a	Proximal femur
PDR28b	Tibial membrane
PDR28c	Tissue distal tibia



PDR139a	Capsule
PDR139b	Femoral neck tissue
PDR139c	Superficial hip fascia

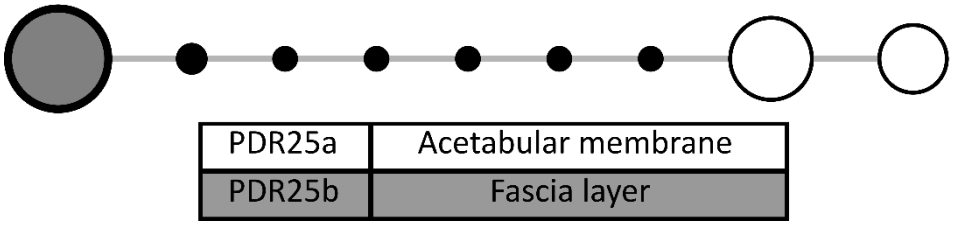


TF16a	Deep tissue
TF16b	Deep tissue
TF16c	Deep tissue

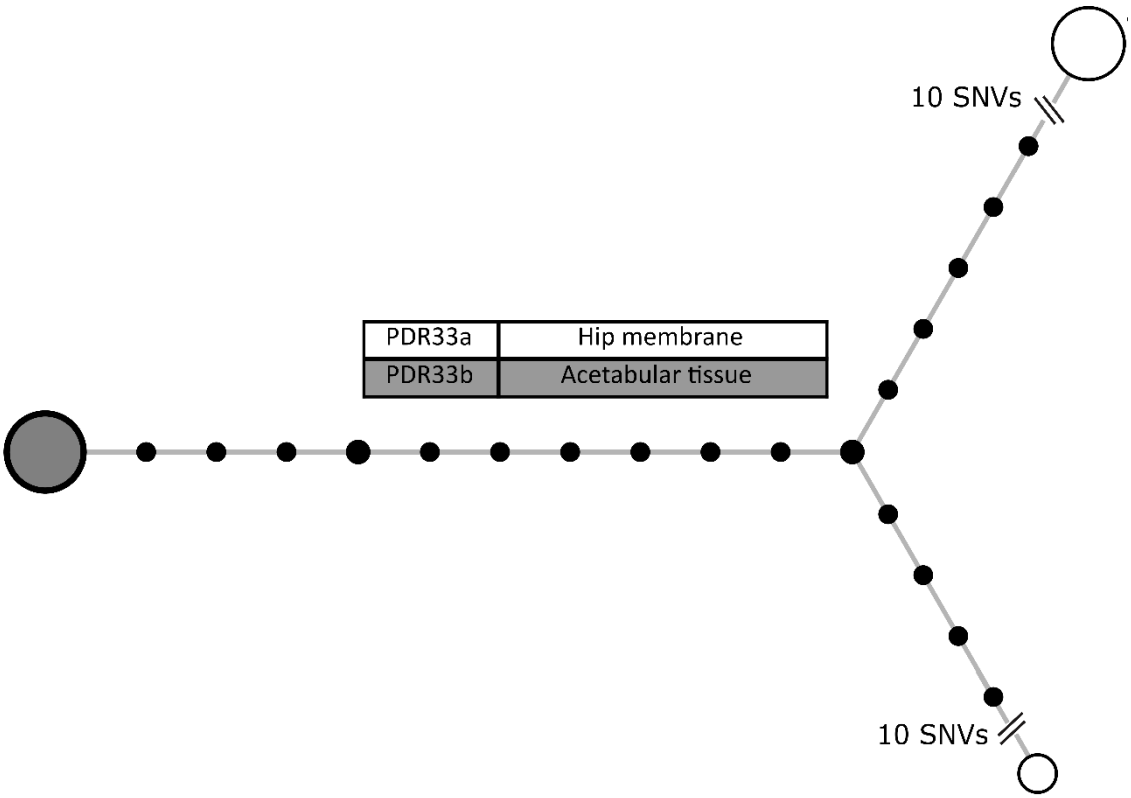


TF26a	Tissue
TF26b	Tissue
TF26c	Tissue

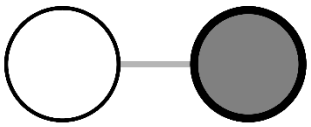
Figure *4.5*c. Haplotrees drawn from the detected genotypes of *S. epidermidis* colonies cultured from ODRI. The size of each node is proportional to the number of times a genotype was detected. Black nodes represent unobserved intermediate hypothetical genotypes. Grey edges represent the mutations that separate genotypes White/grey shading indicates specimen.



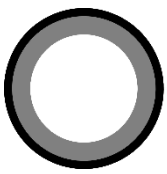
PDR25a	Acetabular membrane
PDR25b	Fascia layer



PDR33a	Hip membrane
PDR33b	Acetabular tissue



PDR36a	Joint fluid
PDR36b	Femoral membrane



PDR39a	Knee fluid
PDR39b	Tibal tunnel tissue

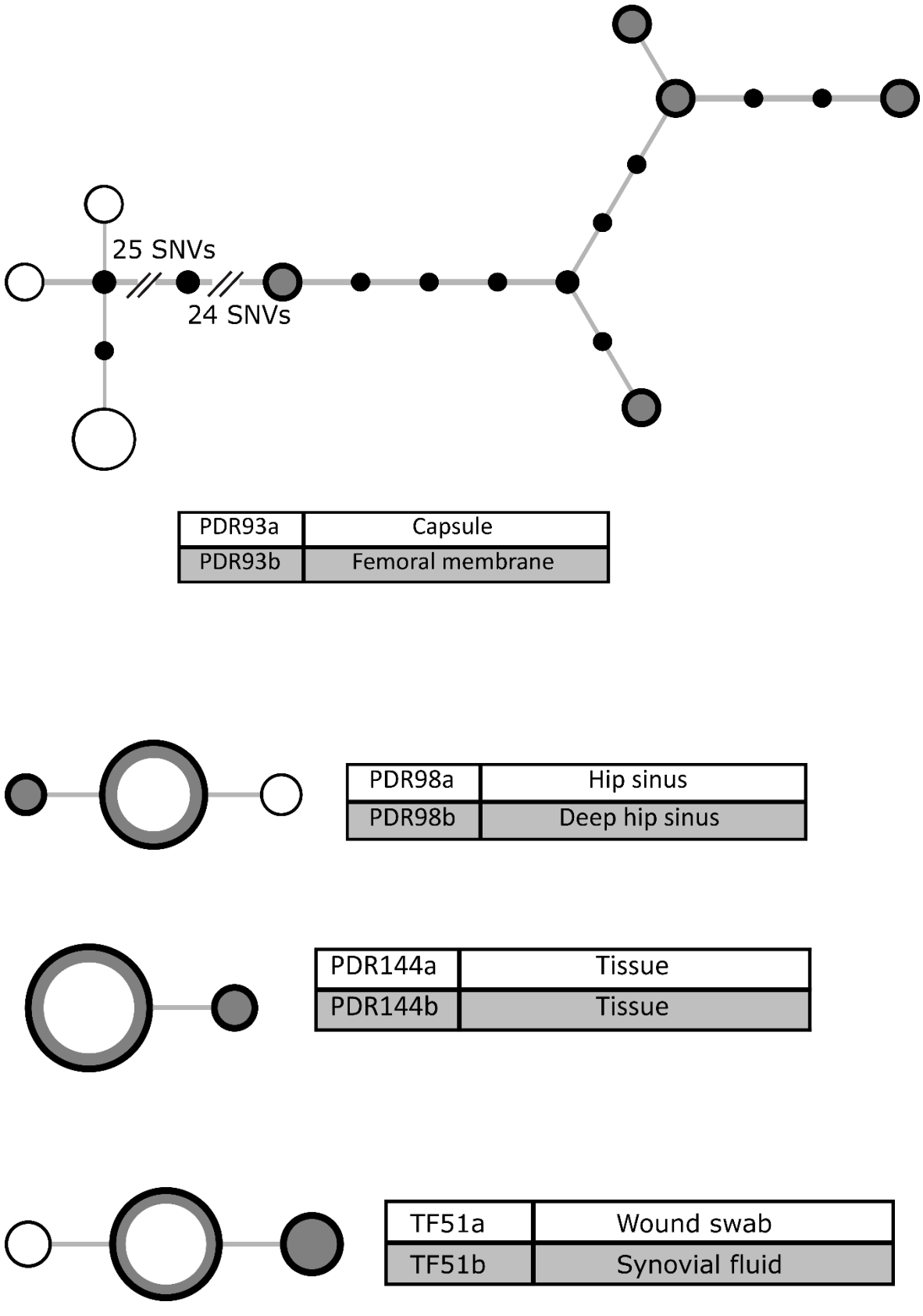


Figure *4.5*d. Haplotrees drawn from the detected genotypes of *S. epidermidis* colonies cultured from ODRI. The size of each node is proportional to the number of times a genotype was detected. Black nodes represent unobserved intermediate hypothetical genotypes. Grey edges represent the mutations that separate genotypes White/grey shading indicates specimen.

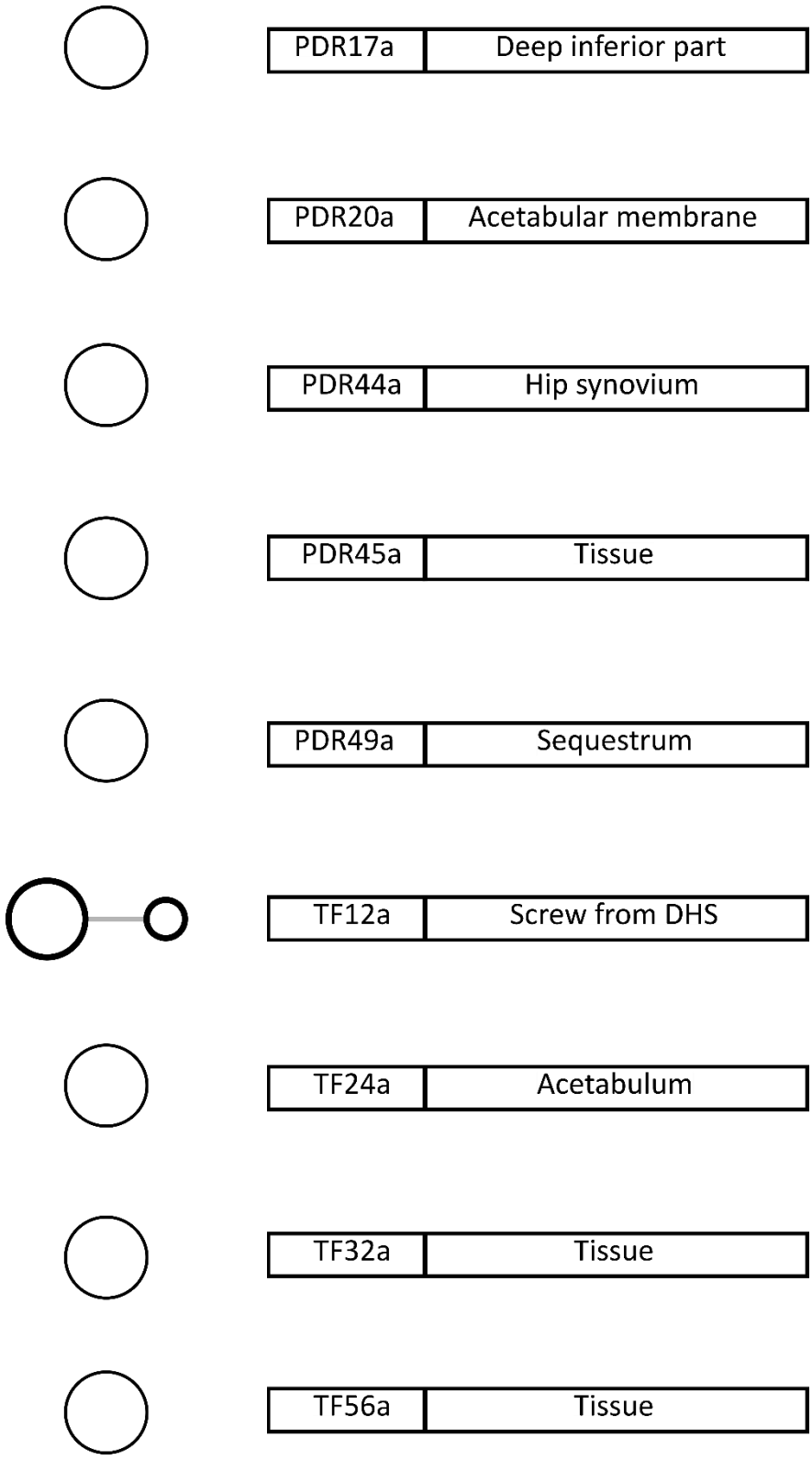


Figure *4.5*e. Haplotrees drawn from the detected genotypes of *S. epidermidis* colonies cultured from ODRI. The size of each node is proportional to the number of times a genotype was detected. Grey edges represent the mutations that separate genotypes

4.4.3 Determining an average SNV rate for cases of ODRI

Orthopaedic device-related infection events are, in effect, transmission events, in which a bottleneck occurs where bacterial diversity is greatly reduced, because of the small numbers of bacterial cells which comprise the inoculum, and because of barriers to infection such as the host immune response and antimicrobial prophylaxis, or because of selective sweeps. This can result in a founding population that has very low genomic diversity. Applying the assumption that a complete genomic bottleneck event (involving a single genotype) has occurred, it may be possible to use the time from device implantation to culture to determine the average rate of SNV accumulation within the infecting population. This number would represent the slowest possible rate of SNV accumulation since it could not account for unobserved bottleneck events that might also occur after device implantation. Therefore, the maximum expected rate of diversification from the most recent common ancestor to time of culture could be calculated as $2m$ where m represents the genomic mutation rate.

Figure 4.6 shows an attempt to apply this method to the cases of ODRI in this study where the time of implantation was known. The exact date of implantation was not known for Case PDR9, only the year. Therefore, three different results representing the minimum, maximum and median time from implantation to culture generated a boxplot. Three other cases (PDR1, PDR93 and TF26) also had approximate implantation dates however since only the month was known the timeframes were too narrow to generate a visible boxplot. As described previously, case PDR38 cultured two different sequence types of *S. epidermidis* with very little or no diversity between genomes of the same ST. Therefore, the two STs isolated from PDR38 were analysed separately. Figure 4.6 shows that during the cases of ODRI $2m$ ranged from 0 to 93.14 per year. The mean $2m$ was 14.38 SNVs per year and the median was 4.94 SNVs per year. This data is likely to be affected by the availability of resources required for population growth and cell death because of antimicrobial therapy and the host immune system. Moreover, it is not known whether the initial inoculation event was a complete bottleneck, for example PDR38 had two infecting sequence types, one of which showed a pairwise distance of 93.14 SNVs after just 19 days suggesting a gross inoculation of the implanted device.

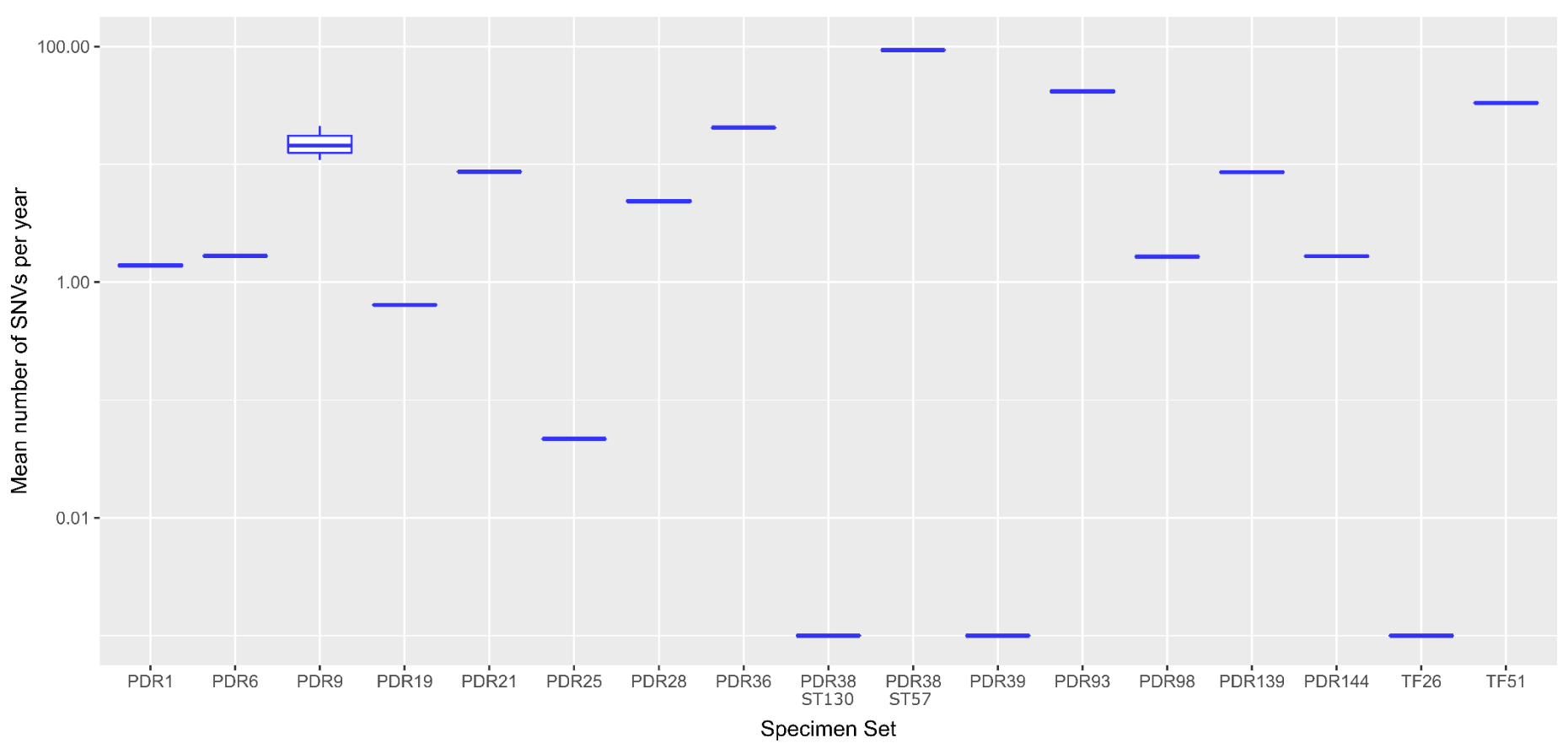


Figure 4.6. Box plots of mean number of SNVs per year of each case of ODRI. Box plot of PDR9 represents an unknown implantation date (sometime in 2015) generating a range of results.

4.4.4 Linear regression of the increase in diversity of *S. epidermidis* infection over time.

A linear regression analysis was performed to determine the accumulation of diversity over time. Cases that cultured more than one sequence type were removed from the analysis. Six cases in which the date of organism isolation was more than 24 months (730 days) after surgical implantation were removed from the analysis as these were most likely associated with haematogenous seeding (356) from another site and so it was not possible to determine the date of device inoculation. In addition, it was not possible to determine the exact date of device implantation for four of the cases of ODRI from the clinical history however timeframes were known and so used in the analysis with the median date used as an estimate. Table 4.10 shows that for the ODRI specimens, the average number of SNVs accumulated between device implantation and culture was 7.6 (95% CI 1.43 – 13.8) occurring over an average of 223.2 days which equates to an average of 1 SNV across the entire genome every 29.4 days.

Samples	Parameter	Mean	Standard Deviation
ODRI	SNVs	7.6	12.2
	Days	223.2	256.6
Contaminants	SNVs	0	0
	Days	230.6	188.3

Table 4.10. Summary of the number of days and SNVs between time of device implantation and specimen culture.

Figure 4.7 shows the mean pairwise distance of the genomes against time from device implantation to organism isolation for the cases of ODRI and specimen contamination. The regression line intercept was set at 0 with the assumption that the inoculation of either the implanted device or contaminated specimen was monoclonal and the result of a bottleneck event in which there was a single most recent common ancestor.

The regression line of ODRI shows an increase in the mean number of SNVs over time however the co-efficient of determination value, R^2, was very low at 6.57% meaning there was very little correlation between time and the accumulation of diversity in these isolate genomes. The correlation co-efficient was 0.54 suggesting that variations in time from implant to culture had a moderate effect on variations in the accumulation of diversity. Conversely, the contaminant isolates show no discernible increase in diversity regardless of the duration between implantation and isolation.

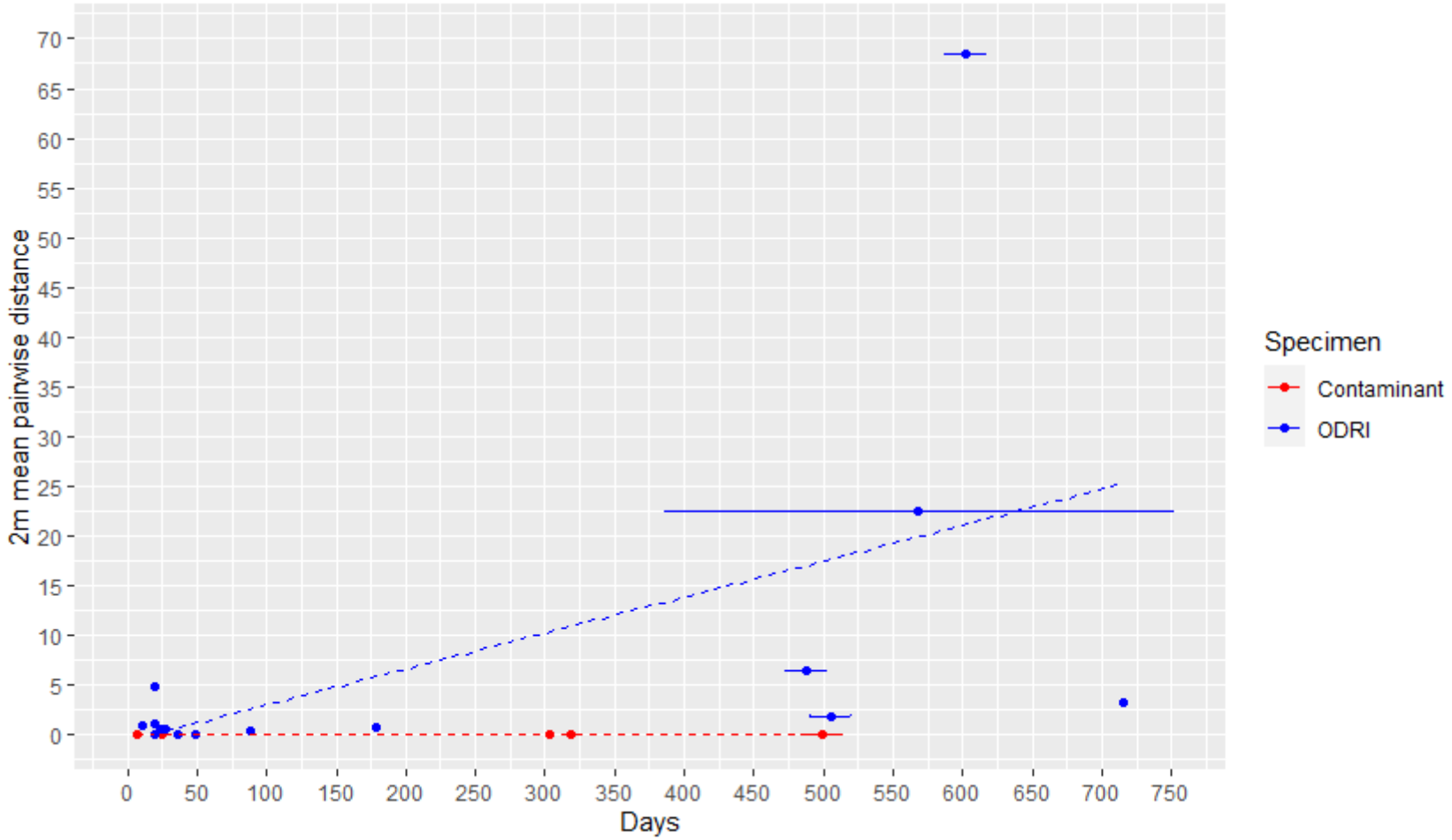


Figure 4.7. Plot of the mean pairwise distance of genomes of *S. epidermidis* from orthopaedic device-related infection-associated and contamination-associated cultures against the time from device implantation to culture of peri-prosthetic specimens. Where the exact date of implantation was not known horizontal lines indicate the known timeframe and the dot indicates the middle date. Specimens with a 2*m* pairwise distance greater than 70 SNVs were excluded as they were a result of polyclonal infection.

4.5 Discussion

The ability to accurately differentiate between infective and non-infective causes of orthopaedic device implantation failure is essential to guide patient management and to treat infection. A significant body of work has already been carried out to improve and consolidate methods for the diagnosis of ODRI, including the establishment of algorithms employing clinical and laboratory criteria (76,265). The application of these algorithms may be more problematical when the infection is sub-acute and the organism a constituent of the normal flora, such as *S. epidermidis*. Several studies have attempted to address the challenges of determining whether a cultured *S. epidermidis* should be regarded as the agent of a true infection or as contaminant. For example, biological factors have been examined that might predispose an organism to invasiveness (32,97,105,109,110,144). The work outlined in this chapter has attempted to determine if genomic characteristics of cultured isolates can be used to discriminate between infecting- and non-infecting isolates. Within-host evolution of pathogens is a significant phenomenon whereby mutations are selected that allow organisms to undergo adaptive change to a new environment and respond to the host immune system and antibiotic exposure. In addition to adaptive evolution, the normal substitution rate (molecular clock) may drive diversity of an infecting population, as has been observed during carriage (318). Whole-genome sequencing has been employed to characterise and measure within-host population dynamics including the accumulation of diversity following the bottleneck events of transmission and invasion of *S. aureus* (285,308,318). Similar methods have been employed in this study to attempt to characterise differences between infecting and contaminant populations of *S. epidermidis*.

The diversity amongst the *S. epidermidis* isolates analysed in this study was predictably lower than that seen in the normal human carriage isolates of the previous chapter. In total there were 18 sequence types detected amongst the 65 patient cases (contrasting with the 32 STs detected amongst five people of Chapter 3.). Of interest, however, was the detection of six (6/18) novel STs. The novel ST781 was one allele different (it was completely lacking the *arcC* gene) from ST2. Furthermore, amongst the eight ST2 strains isolated, eight different phenotypes were observed, suggesting that success of this lineage may be associated with considerable phenotypic diversity. Further investigation showed that despite having differing sequence types and antibiograms, ST2 and ST781 were markedly similar genomically and so were treated as a common ST in downstream analyses. Two other cases yielded different strains according to either MLST or antibiogram. Case PDR38 cultured ST57 and ST130 (with different

antibiograms) from multiple specimens each and was thus defined as a 'true' case of polyclonal infection. PDR93 cultured *S. epidermidis* with two different antibiograms. However, both populations were ST2 and were closely related (with a minimum of 51 SNVs), which suggests a recent common ancestor. Furthermore, only 2/5 of the positive cultures from the patient were sequenced. If the other three had been investigated, they may perhaps have yielded intermediate genotypes. In the three cultures investigated for case TF26, two different sequence types and antibiograms were observed, which may have been the result of polyclonal infection or contamination by ST558. These doubts might have been resolved by investigating more isolates. It may be possible to define a threshold of relatedness, as has been proposed in Chapter 3, for commensal carriage populations. Indeed, this may in fact be simpler since there should be little or no admixture of mobile genetic elements between lineages that could potentially confound the data.

Sequencing individual colonies from multiple cultured specimens yielded a higher pairwise distance than the pairwise distance of colonies from individual cultures. This resulted in a higher correlation with the clinical diagnosis for multiple cultures ($p = 0.000115$) than for single cultures ($p = 0.0254$). The study showed that a low pairwise distance of zero or one was usually associated with contamination and could also be associated with infection, but that a higher pairwise distance was always associated with infection. It is possible that sequencing more colonies from the same specimen or sequencing colonies from more specimens from the same site of infection would show more diversity and therefore provide clearer evidence for infection. However, this may not be the case as external factors such as selective sweeps or bottlenecks caused by the host immunity or antibiotic treatment may reduce the bacterial population diversity. Furthermore, whilst polyclonal contamination was not observed during this study it can and does occur and so must be considered. Polyclonal contamination may be readily detected if criteria are established for differentiating between strains. For example, the previous chapter reported a positive predictive value of 100 SNVs for differentiating subtypes of commensal *S. epidermidis* with 95% confidence. However, since the macro-diversity of clinical strains of *S. epidermidis* is smaller than the macro-diversity of commensal strains, it may be that the micro-diversity of clinical strains is also smaller and therefore a PPV of less than 100 SNVs may be determined. Explicit characterisation of polyclonal contamination of specimens and cultures could therefore be a valuable further investigation.

The production of haplotrees to visualise the population structure of the infecting organism could be a better aid to diagnosis than simply a measure of pairwise distance. Some of the haplotrees showed marked radial structures which may represent how the infecting population

has diversified. By inserting an outgroup of a genome of a closely related but different strain into the tree structure it may be possible to identify the ancestral genotype of the sequenced colony genomes. Therefore, building up a larger library of genomes from colonies of *S. epidermidis* associated with ODRI may add further value to the method.

Determining a molecular clock for infecting populations of *S. epidermidis* was less successful than discriminating between infecting and contaminating populations. This was likely due to several factors including a small sample size, unknown specific dates of orthopaedic device implantation and the influence of exogenous environmental factors on bacterial population - and therefore diversity - size. Exogenous factors such as host immunity or antibiotic insult may cause a reduction or slower increase in bacterial population diversity. The formation of a bacterial biofilm is considered a reaction to the external environment which results in the deceleration of cell growth (77). This slower growth will, in turn, slow the normal accumulation of population diversity and therefore affect any measurement of a molecular clock. Similarly, a successful selective sweep may reduce the diversity of the population, returning it back to a level like a previous point in time. However, since a population diversity of greater than 1 SNV was statistically more likely to originate from ODRI then even if the diversity is reduced or only increases slowly *in vivo* whilst attempting to identify correlation between time and diversity may not be a useful tool, identifying any diversity still is useful.

As a potential tool to aid in the diagnosis and management of ODRI, characterising the diversity of the bacterial population would perhaps be incorporated into the existing well-established criteria of laboratory and clinical observations (76,265). At present, sequencing multiple colonies from multiple cultured specimens has greater time and cost implications however given the high cost and time already required to diagnose and treat implant failures this may not be proportionally high, especially if it increases the chance of successful treatment. WGS data derived from infecting populations would also yield useful information on genetic determinants of antimicrobial resistance (as discussed in Chapter 5) and virulence factors which may help in the management of the infection. Furthermore, building up a database of the strains isolated from each patient could yield valuable information on relatedness which could help identify epidemic or endemic strains at a higher resolution afforded by multi-locus sequence typing. A potential area of further research would be to try and build on the work of metagenomic sequencing of intra-operative specimens (341,342) by incorporating a similar a SNV-based method of characterising the infecting population.

4.6 Conclusion

In this chapter I have demonstrated that there is a difference between the pairwise distance of genomes of *S. epidermidis* isolated from cases of orthopaedic device-related infection and from cases of contamination.

Measuring the pairwise distance of genomes of isolates associated with infection and contamination yielded statistically significant results which could be exploited to aid diagnosis. Since a pairwise distance greater than one SNV was always associated with ODRI this could be included in the criteria for discriminating between infection and contamination.

The generation of haplotrees could potentially be used to visualise infecting populations which could help characterise ODRI at the genomic level and could be used alongside existing clinical and laboratory methods to characterise the infection.

The attempt to determine a correlation between time from implantation to culture and pairwise diversity did not produce statistically significant results.

Chapter 5
Detecting genomic determinants of antimicrobial resistance in *Staphylococcus epidermidis*

5.1 – Introduction

Staphylococcus epidermidis can colonise different ecological niches including, humans and other animals, plants and the environment (260,357–361). The ability of *S. epidermidis* to survive in a range of habitats is facilitated by a flexible genome and the acquisition of genes that confer resistance to environmental pressures (262,359). Nosocomial strains of *S. epidermidis* which have become adapted to the hospital environment may become endemic or associated with outbreaks of infection in specialist units (10,238,255,362). This is, in part, mediated by the ability to resist antimicrobial compounds. Antibiotic resistant strains have been isolated from patients and staff (where they may live as commensals), environmental samples and from sites of infection (29,105,299,363). The clonal expansion of antibiotic resistant sequence types such as ST2 and ST27 within healthcare environments attests to the selective advantage conferred by antibiotic resistance in this environment (238,255). Moreover, these successful strains continue to accumulate genetic factors which confer resistance to antibiotics and are disseminated in healthcare settings across the world (10,362,364).

S. epidermidis orthopaedic device-related infections are often difficult to treat because of antibiotic resistance and the ability of the organism to form biofilms on biosynthetic materials. Knowledge of the isolate's antibiotic susceptibility is necessary as the number of compounds that can efficaciously penetrate to the site of infection and biofilm is limited.

Traditional phenotypic methods of determining resistance or susceptibility rely on determining if the antibiotic prevents the isolate growing *in vitro*. Molecular methods have been devised that can detect the genes or point mutations associated with resistance. Multiplex polymerase chain reaction (PCR) and DNA microarray assays have been developed that can detect the antibiotic resistance genes found in clinically important isolates of staphylococci (365–367). More recently, tools have been developed that can detect antibiotic resistance-associated genes and point mutations from whole genome sequences. These tools include ResFinder, PointFinder, CARD (Comprehensive Antibiotic Resistance Database), ARIBA (Antimicrobial Resistance Identification By Assembly), Mykrobe and ARG-annot (Antibiotic Resistance Gene-ANNOTation) and BLAST (286,368–373). Moreover, a number of studies have compared phenotypic methods with WGS-based methods to successfully predict antibiotic resistance for a number of pathogens including *S. aureus*, *Mycobacterium tuberculosis*, *Escherichia coli*, *Klebsiella pneumoniae*, *Neisseria gonorrhoea, Streptococcus pneumoniae, Helicobacter pylori* and *Shigella sonnei* (204,374–379). As well as predicting resistance, some studies have used WGS data to

predict the actual minimum inhibitory concentrations of some organisms to generate a quantitative result (166,375,380–382).

It is therefore plausible that one or more of the listed methods could be applied to whole-genome sequences from *S. epidermidis* to infer an antibiotic resistance profile.

5.2 - Objectives

This chapter outlines the methods and results of developing a BLASTn-based method of detecting genes and mutations associated with *S. epidermidis* resistance to a panel of antibiotics routinely tested at the Royal Sussex County Hospital Microbiology Department. This builds on previous research by Gordon *et al.* (204) which used BLASTn to predict antimicrobial resistance in *S. aureus*. The work outlined in this chapter had four components:

1) Perform phenotypic antimicrobial susceptibility testing on the collection of *S. epidermidis* isolates from normal human carriage and orthopaedic device-related infection,
2) Create a catalogue of the gene sequences of genetic determinants of resistance to a panel of 13 antibiotics commonly used to manage or prevent infection,
3) Develop and pilot a BLASTn-based method to interrogate the *S. epidermidis* genomes for the resistance conferring gene sequences,
4) Describe levels of concordance between the genotypic and phenotypic methods.

5.3 Methods

5.3.1 Sampling

Two collections of isolates of *S. epidermidis* were investigated in this study. One was the collection of isolates from healthy volunteers as described in Chapter 3, Section 3.3.1 which were obtained with ethical approval by Brighton and Sussex Medical School Research Governance and Ethics. The other was the collection of clinical isolates obtained from cultured specimens relating to suspected orthopaedic device-related infections as detailed in Chapter 4, Section 4.3.1 and did not require ethical approval under the Statutory Instrument Regulations 2002 No. 1438 (348). The use of patient or participant data was not required to conduct this study.

5.3.2 Culture and Identification

5.3.2.1 Healthy carriage isolates

As previously described, swabs from seven anatomical sites (anterior nares, axillae, cheeks and throat) were inoculated in a 10ml 7% HPA Salt Broth (Oxoid Ltd, Basingstoke, UK) (307). The salt broths were incubated overnight at 35°C. A 10µl loopful of the inoculated broth was spread onto a Columbia CAP Staph/Strep Selective Agar & Horse Blood (Oxoid Ltd, Basingstoke, UK) which were incubated aerobically at 35°C for up to 48 hours. Colonies of suspected *S. epidermidis* were initially selected according to their visual appearance; small, white colonies on Columbia Cap Selective Agar (Oxoid Ltd, Basingstoke, UK) and sub-cultured on to a Columbia Blood Agar (CBA) plate (Oxoid Ltd, Basingstoke, UK) and incubated aerobically at 35°C overnight. Following incubation, colonies were identified to species using MALDI-ToF MS (Bruker Daltonik GmbH, Leipzig, Germany), (see Chapter 2: Materials and Methods).

5.3.2.2 Clinical isolates

As described in Chapter 4, Section 4.3.2, specimens from suspected cases of ODRI were investigated at the John Radcliffe Hospital microbiology laboratory or the Royal Sussex County Hospital microbiology laboratory as part of routine clinical practice. Cultures of *S. epidermidis* were investigated as part of this study. Five colonies from original culture plates of Columbia Blood Agar or Chocolate Agar (Oxoid Ltd, Basingstoke, UK) were harvested to capture the diversity of cultured organism. Isolates from suspected ODRI and suspected contamination underwent examination for this study. Colonies were confirmed as *S. epidermidis* using MALDI-ToF MS (Bruker Daltonik GmbH, Leipzig, Germany), (see Chapter 2: Materials and Methods).

5.3.3 Antimicrobial Susceptibility Testing

Antimicrobial susceptibility testing was performed on each isolate. In the first instance the disc diffusion method was employed for testing cefoxitin, penicillin, erythromycin, gentamicin, fusidic acid, clindamycin, tetracycline, ciprofloxacin, linezolid, rifampicin and mupirocin and any discrepant results were then checked by using MIC strips, where available. The MIC strip method was used for susceptibility testing to teicoplanin and vancomycin in the first instance as the disc diffusion method is considered to be an unreliable indicator of susceptibility (279).

5.3.3.1 Disc diffusion

The disc diffusion method is outlined in Chapter 2: Materials and Methods (279). Briefly, three to five colonies of similar appearance were lightly touched with a 1µl loop. This was inoculated into a vial of 5ml sterile 0.9% saline (Oxoid Ltd, Basingstoke, UK) to achieve a 0.5 McFarland suspension.. A sterile cotton-tipped applicator was used to spread the inoculum onto an Iso-Sensitest Agar plate (Oxoid Ltd, Basingstoke, UK). Antimicrobial Susceptibility Discs (Oxoid Ltd, Basingstoke, UK) were applied to the inoculated plate. Plates were then incubated 18 – 20 hours aerobically at 35°C. Zones of inhibited growth around the antibiotic disc were measured to determine resistance (R), intermediate category (I) or susceptibility (S) according to the clinical breakpoints outlined by BSAC and as shown in Table 2.1 (280).

5.3.3.2 Minimum inhibitory concentration

A 0.5 McFarland standard of inoculum was prepared in sterile saline as above. This was used to inoculate an Iso-Sensitest Agar plate (Oxoid Ltd, Basingstoke, UK). An E-test© strip (BioMerieux, Marcy-l'Étoile, France) or M.I.C. Evaluator strip (Oxoid Ltd, Basingstoke, UK) was then applied to the centre of the plate. Plates were incubated 18 – 20 hours aerobically at 35°C. For oxacillin testing, Columbia Agar Base with 2% Salt (CSA) plates (Oxoid Ltd, Basingstoke, UK) were used as the growth medium and plates were incubated at 30°C for 24 hours before reading. Following incubation, MICs were interpreted as the lowest concentration of antibiotic that inhibited the growth of test organism as shown in Table 2.2.

5.3.3.3 Nitrocefin disc test for beta-lactamase

Nitrocefin is a chromogenic beta-lactam susceptible to all lactamase enzymes. The nitrocefin disc test (Mast Group Ltd., Bootle, UK) was used to detect the production of penicillinase.

5.3.4 Selection of Isolates for Genotyping

All isolates described in Chapter 3 and Chapter 4 underwent phenotypic antimicrobial susceptibility testing. However, many isolates in the two collections were either genomically identical or very closely related. Figure 5.1 outlines the criteria used for the selection of isolates for genotyping to avoid duplication of results.

Two isolates from the same sequence type (ST2) and from the same patient (PDR25) were included in the analysis because although both were phenotypically resistant to cefoxitin by disc

diffusion method one had a zone of inhibition measuring 0 mm and the other had a zone of inhibition of 19 mm.

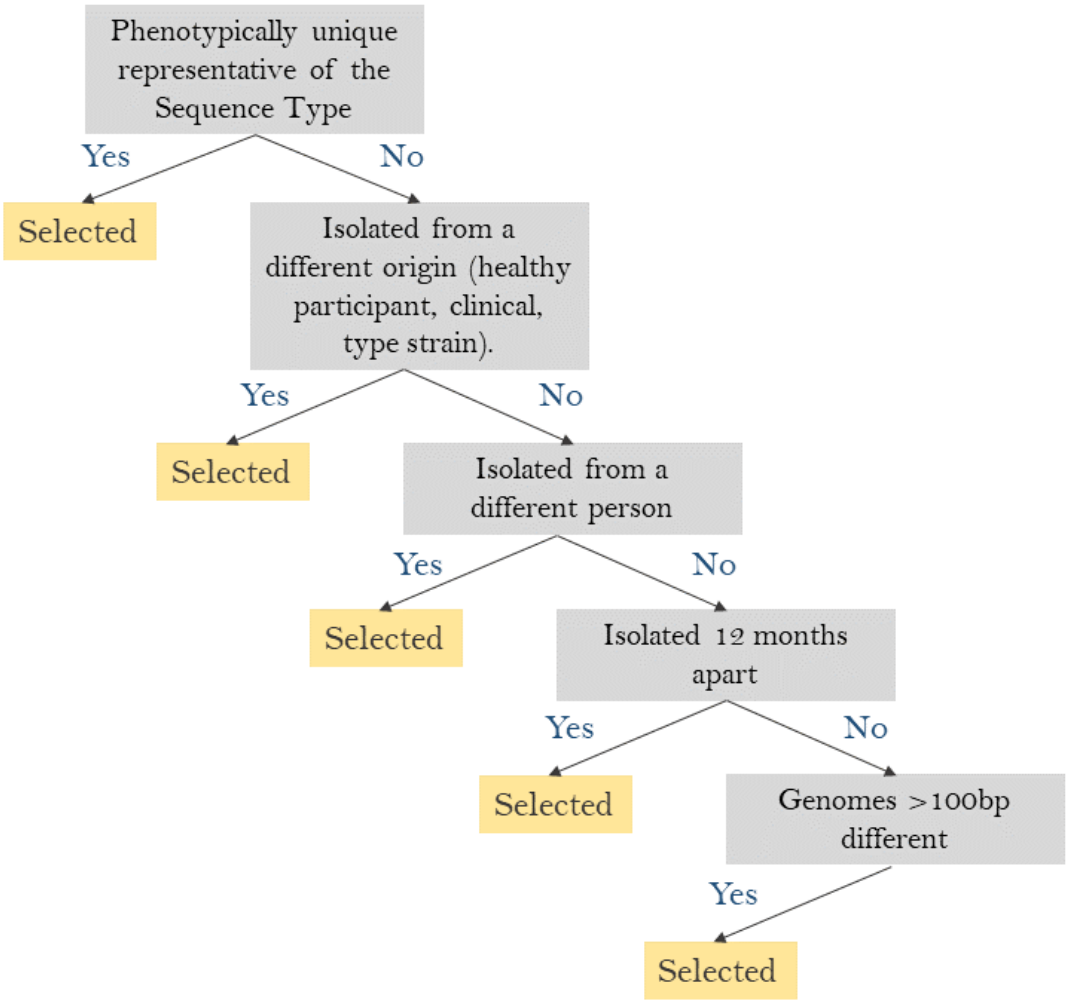


Figure 5.1. Criteria for selecting *S. epidermidis* isolates for detection of genetic determinants of antimicrobial resistance.

5.3.5 Detection of Genetic Elements Associated with Antimicrobial Resistance

5.3.5.1 Literature search for common determinants of antimicrobial resistance

The method developed by Gordon *et al.* (2014) was used to detect genetic elements conferring antimicrobial resistance in the *S. epidermidis* isolates (204). A Google Scholar search using the headings "Staphylococcus epidermidis" and "resistance" along with the antibiotics listed in

Table 2.2 was used to identify the common genetic causes of resistance in *S. epidermidis*. The search was then broadened by replacing "Staphylococcus epidermidis" with "Staphylococcus" or "staphylococci". This search yielded two datasets; acquired resistance genes and mutations that confer resistance in antibiotic targets.

5.3.5.2 Acquired genes associated with resistance

Once a list of the commonly acquired resistance genes was made, sequences obtained from either *S. epidermidis,* or other staphylococci were compiled from the NCBI Genbank resource into a single FASTA file. A summary of the details of these sequences is outlined in Table 5.1.

5.3.5.3 Amino acid substitutions, insertions or deletions in antimicrobial targets that are associated with resistance

A list of antimicrobial targets along with and amino acid variations that are associated with antimicrobial resistance was collected (as listed in Table 5.2). The amino acid sequences of the target proteins were obtained from translated gene sequences from the NCBI Genbank resource and compiled into a single FASTA file.

5.3.5.4 Use of BLAST to detect genetic determinant of resistance

The datasets were then used as query sequences in BLAST searches of the study genomes. Details of these methods are summarised in Table 5.3. BLASTn was used to detect the acquired gene sequences listed in Table 5.1. and output relative coverage scores (the length of the sequence that has aligned multiplied by the percentage of identical nucleotide matches). A threshold of >80% relative coverage was applied as a filter to remove low-scoring nucleotide alignments. tBLASTn was used to align the amino acid sequences listed in Table 5.2 in the *S. epidermidis* genomes and identify any amino substitutions, insertions or deletions. The detection of amino acid changes listed in Table 5.2 was recorded.

5.3.5.5 Use of Clustal W to align sequences

The Clustal W programme aligns multiple DNA sequences (383). The Clustal W (Codons) version translates the codons to amino acids, performs the alignment and then replaces the amino acids

with the original codons (383). Clustal W was used to investigate resistance-associated genes for the presence of non-synonymous mutations or insertions or deletions.

Antimicrobial	Gene	Product	Genbank Accession	Species	Isolate Identifier
Ciprofloxacin	*norA*	Efflux pump	NC_006663.1	*S. epidermidis*	RP62A
Clindamycin and/or Erythromycin	*ermA*	rRNA adenine N-6-methyltransferase	NC_006663.1	*S. epidermidis*	RP62A
	ermC	rRNA adenine N-6-methyltransferase	HG970733.1	*S. epidermidis*	C6187
	ermT	rRNA adenine N-6-methyltransferase	KX274135.1	*S. arlettae*	SA-01
	lnuA	Nucleotidyltransferase	AM184101.1	*S. haemolyticus*	coa101
	mphC	Phosphotransferase	AM397632.1	*S. equorum*	CSNO44
	msrA	Efflux pump	EF092840.1	*S. epidermidis*	Unpublished
Fusidic Acid	*fusB*	Ribosomal protection protein	NG_047902.1	*S. epidermidis*	NTUH-2793
	fusC	Ribosomal protection protein	CP033782.1	*S. epidermidis*	FDAARGOS_529
Gentamicin	*ant(6)-I*	Adenylyltransferase	NC_006663.1	*S. epidermidis*	RP62A
	ant(9)-I	Adenylyltransferase	NC_006663.1	*S. epidermidis*	RP62A
	aph(2'')	Phosphotransferase	NC_006663.1	*S. epidermidis*	RP62A
	aph(3')-IIIa	Phosphotransferase	NC_006663.1	*S. epidermidis*	RP62A
Linezolid	*cfr*	RNA large subunit methyltransferase	JX910899.1	*S. epidermidis*	426-3147L
Methicillin	*mecA*	Penicillin binding protein	KF415244.1	*S. epidermidis*	TN/CN/KSE/1/13
	mecC	Penicillin binding protein	KT192641.1	*S. aureus*	TRN6234
Mupirocin	*mupA*	Isoleucyl-tRNA synthetase	JTAY02000002.1	*S. epidermidis*	FDAARGOS_83
Penicillin	*blaZ*	Beta-lactamase	NC_006663.1	*S. epidermidis*	RP62A
Tetracycline	*tetK*	Efflux pump	176280.1	*S. epidermidis*	ATCC 12228
	tetL	Efflux pump	HG380319.1	*S. epidermidis*	47
	tetM	Ribosomal protection protein	HG380319.1	*S. epidermidis*	47
Vancomycin and/or Teicoplanin	*vanA*	Dipeptide terminus modifying protein	MG592387	*S. aureus*	Transposon Tn1546

Table 5.1. List of acquired resistance genes detected in *S. epidermidis* and other staphylococci obtained from NCBI Genbank.

Antimicrobial	Gene	Product	Genbank Accession	Isolate	Amino Acid Change
Ciprofloxacin	*grlA*	DNA topoisomerase	ASJ93525.1	ATCC 12228	V41G, I45M, A48T, D79V, S80F, S80Y, V82L, Y83N, E84K, E84G, E84V, D84A, D84C, S108N, A116E, D432G
	grlB	DNA topoisomerase	ASJ93526.1	ATCC 12228	E422D, P585S, D443E, R444S, P451S, R470D, N404S
	gyrA	DNA gyrase	ASJ94484.1	ATCC 12228	S84F, S84L, S84Y, E88K, S85P, E88G, E88L, G106D
Fusidic Acid	*fusA*	Translocase elongation factor	ASJ94281.1	ATCC 12228	A67T, A70V, A71V, R76C, V90A, V90I, P114H, Q115L, A160V, M161I, D189G, D189V, E233Q, D373N, A376V, T385N, T387I, P404L, P404Q, P406L, S416F, L430S, B434N, T436I, H438N, F441Y, E444K, E444V, E449K, G451V, G452C, G452S, M453I, L456F, H457Q, H457Y, L461K, L461S, D463G, R464C, R464H, R464S, C473S, P478S, G556S, V607I, G617D, M651I, F652S, Y654N, A655E, A655P, A655V, T656K, R659C, R659H, R659L, R659S, G664S
	fusE	Ribosomal protein	ASJ92851.1	ATCC 12228	Q140L
Linezolid	*rplC*	Ribosomal protein	ASJ92836.1	ATCC 12228	G355A, G463X, A505T, H146Q, V154L, A157R, G137S, M156T, M169L
	rplD	Ribosomal protein	ASJ92837.1	ATCC 12228	A202C, K68N, L108S, ins71GR, ins71GGR
	rplV	Ribosomal protein	ASJ92841.1	ATCC 12228	A138G, C141T, G166A
Mupirocin	*ileS*	Isoleucyl-tRNA synthetase	NP764423.1	ATCC 12228	V588F, V631F
Rifampicin	*rpoB*	RNA polymerase	ASJ94286.1	ATCC 12228	Q456K, S463P, S464P, L466S, Q468K, Q468L, Q468R, M470T, D471E, D471G, D471Y, A473T, N474K, ins475H, insG475, A477D, A477T, A477V, H481D, H481N, H481Y, R484H, S486L, S486F, I527F, I527L, I527M, S529L, A534V, D550G, Q565R

Table 5.2. List of housekeeping genes present in *S. epidermidis* obtained from NCBI Genbank, and amino acid substitutions, insertions or deletions associated with antimicrobial resistance.

Query Sequences	Query Sequence Molecules	Subject Sequences	Subject Sequence Molecules	Programme	Indicator of Genotypic Resistance
Acquired resistance genes	Nucleotides	*S. epidermidis* genomes	Nucleotides	BLASTn	Relative coverage
House-keeping genes	Amino acids	*S. epidermidis* genomes	Nucleotides	tBLASTn	Amino acid changes

Table 5.3. Summary of BLAST methods employed for interrogating *S. epidermidis* genomes for antimicrobial resistance determinants.

5.4 Results

5.4.1 MLST and Initial Phenotyping

The selection of isolates, based on the criteria outlined in Section 5.3 Methods, yielded 87 isolates. 48/87 (55%) isolates were selected from healthy individuals as discussed in Chapter 3, 38/87 (44%) isolates were selected from clinical cases of suspected orthopaedic device-related infection as discussed in Chapter 4 and 1/87 (1.1%) was the type strain NCTC 11047. These criteria yielded 45 different sequence types. There were differences between the two collections with the isolates from healthy carriage being more diverse according to sequence type (Table 5.4) but exhibiting less phenotypic resistance to the antimicrobials tested than those from a clinical source (Table 4.6 and Table 4.7).

5.4.1.1 Isolates from healthy individuals

Of the 48 isolates selected from healthy study participants, 22 (45.8%) were unique representatives of a sequence type, the remaining 26/48 (54.2%) represented 10 different sequence types meaning 32 different sequence types were represented in the analysis.

Testing the 48 isolates against the 13 antibiotics listed in Table 2.1 and Table 2.2 yielded 624 phenotypic observations, 544/624 (87.2%) of which were susceptible, 78/624 (12.5%) were resistant and 2/624 (0.3%) were intermediate. Overall, fifteen different phenotypic antimicrobial resistance profiles were observed. Antimicrobial susceptibility testing and MLST results are listed in Table 5.4.

Participant	Colony	Cefox	Pen	Eryth	Gent	Vanc	Teic	Fus	Rif	Tet	Cipro	Clind	Lzd	Mup	MLST
F5	F50AxLm1	S	R	R	S	S	S	S	S	S	S	S	S	S	32
	F50AxRm1	S	R	R	S	S	S	S	S	S	S	S	S	S	766
	F50AxRm5	S	R	S	S	S	S	R	S	S	S	S	S	S	767
	F50ChLm1	S	R	R	S	S	S	S	S	S	S	I	S	S	73
	F50ChLm4	S	R	R	S	S	S	R	S	S	S	R	S	S	774
	F50NaRm1	R	R	S	S	S	S	R	S	S	R	S	S	S	190
	F50NaRm6	S	S	S	S	S	S	S	S	S	S	S	S	S	723
	F512AxLm1	S	R	R	S	S	S	S	S	S	S	S	S	S	73
	F512ChLm1	S	R	S	S	S	S	R	S	S	S	S	S	S	778
	F512NaRm1	S	S	S	S	S	S	S	S	S	S	S	S	S	723
	F512NaRm6	S	R	S	S	S	S	R	S	S	S	S	S	S	59
F7	F70AxLm1	S	S	R	S	S	S	S	S	S	S	S	S	S	297
	F70AxRm4	S	S	S	S	S	S	S	S	S	S	S	S	S	73
	F70ChLm2	S	R	R	S	S	S	R	S	S	S	S	S	R	775
	F70ChLm8	S	R	S	S	S	S	S	S	S	S	S	S	S	776
	F70ChRm1	S	S	R	S	S	S	S	S	S	S	R	S	S	769
	F70NaLm4	S	R	S	S	S	S	S	S	S	S	S	S	S	487
	F70NaLm5	S	R	S	S	S	S	R	S	S	S	S	S	S	543
M4	M40AxLm3	S	R	R	S	S	S	S	S	S	S	S	S	S	384
	M40ChLm5	S	R	R	S	S	S	S	S	S	S	S	S	S	153
	M40NaLm2	S	R	S	S	S	S	S	S	R	S	S	S	S	89
	M40NaRm7	S	R	S	S	S	S	S	S	R	S	S	S	S	777
	M412AxLm1	S	R	R	S	S	S	S	S	S	S	S	S	S	384
	M412ChLm2	S	R	S	S	S	S	S	S	S	S	S	S	S	153
	M412NaRm2	S	R	R	S	S	S	R	S	S	S	S	S	S	7

	M412NaRm4	S	R	S	S	S	S	S	S	R	S	S	S	S	777
M7	M70ChRm4	S	R	S	S	S	S	R	S	S	S	S	S	S	73
	M70NaLm1	S	R	R	S	S	S	R	S	S	S	S	S	S	59
	M712AxLm1	S	R	S	S	S	S	S	S	S	S	S	S	S	773
	M712AxLm2	S	R	S	S	S	S	S	S	S	S	S	S	S	558
	M712AxLm4	S	R	S	S	S	S	S	S	S	S	S	S	S	17
	M712AxLm5	S	S	S	S	S	S	S	S	S	S	S	S	S	768
	M712ChLm1	S	S	R	S	S	S	S	S	S	S	S	S	S	457
	M712ChLm5	S	S	R	S	S	S	S	S	S	S	S	S	S	59
	M712ChRm4	S	R	S	S	S	S	R	S	S	S	S	S	S	73
	M712NaLm3	S	S	S	S	S	S	S	S	S	S	S	S	S	190
	M712NaLm4	S	S	R	S	S	S	S	S	S	S	S	S	S	130
	M712NaRm1	S	S	R	S	S	S	R	S	S	S	S	S	S	765
M8	M80AxLm1	S	R	R	S	S	S	S	S	S	S	S	S	S	87
	M80AxLm3	S	R	R	S	S	S	S	S	R	S	S	S	S	770
	M80ChLm2	S	S	S	S	S	S	S	S	S	S	S	S	S	57
	M812AxLm2	S	S	R	S	S	S	S	S	S	S	S	S	S	297
	M812AxLm3	S	R	S	S	S	S	R	S	S	S	S	S	S	73
	M812AxLm6	S	S	I	S	S	S	S	S	S	S	S	S	S	87
	M812AxRm2	S	R	R	S	S	S	S	S	R	S	S	S	S	770
	M812ChLm1	S	S	S	S	S	S	S	S	S	S	S	S	S	329
	M812ChRm2	S	R	R	S	S	S	S	S	S	S	S	S	S	19
	M812NaLm3	S	S	R	S	S	S	S	S	S	S	S	S	S	59

Table 5.4. Antimicrobial susceptibility testing and MLST results of colonies of *S. epidermidis* from healthy participants. S: Susceptible, R: Resistant, I: Intermediate, Cefox: Cefoxitin, Pen: Penicillin, Eryth: Erythromycin, Gent: Gentamicin, Vanc: Vancomycin, Teic: Teicoplanin, Fus: Fusidic Acid, Rif: Rifampicin, Tet: Tetracycline, Cipro: Ciprofloxacin, Clind: Clindamycin, Lzd: Linezolid, Mup: Mupirocin.

5.4.2 Isolates from clinical sources

Of the 39 clinical isolates, 13(33%) were unique representatives of a sequence type and 26/39 (67%) represented 7 sequence types resulting in 20 different sequence types.

Testing the 39 isolates with the panel of 13 antibiotics listed in Table 2.1 and Table 2.2 yielded 507 phenotypic observations, of which 158/507 (31.2%) showed resistance and 349/507 (68.8%) showed susceptibility to the antimicrobials tested (Table 5.5). These observations yielded 28 different antimicrobial resistance profiles.

Source	Colony	Cefox	Pen	Eryth	Gent	Vanc	Teic	Fus	Rif	Tet	Cipro	Clind	Lzd	Mup	MLST
Contaminant	TF11m1	S	S	S	S	S	S	S	S	S	S	S	S	S	520
Contaminant	PDR17am1	R	R	R	R	S	S	R	S	S	R	R	S	S	2
Contaminant	PDR20am1	S	S	R	S	S	S	S	S	S	S	S	S	S	764
Contaminant	PDR44am1	S	R	R	S	S	S	R	S	S	S	S	S	S	208
Contaminant	PDR45am1	S	R	R	S	S	S	R	S	S	S	S	S	S	7
Contaminant	PDR49am1	S	R	S	S	S	S	S	S	S	S	S	S	S	89
Contaminant	TF12m1	S	R	S	S	S	S	R	S	S	S	S	S	S	208
Contaminant	TF24m1	R	R	R	R	S	S	S	S	S	R	R	S	S	2
Contaminant	TF32m1	S	R	S	S	S	S	S	S	S	S	S	S	S	88
Contaminant	TF8m1	S	R	R	S	S	S	S	S	S	S	S	S	S	87
ODRI	PDR12am1	S	R	R	S	S	S	R	S	S	R	R	S	S	2
ODRI	PDR12em1	S	R	R	S	S	S	S	S	S	R	R	S	S	2
ODRI	PDR139am1	R	R	R	R	S	S	R	S	R	R	R	S	S	5
ODRI	PDR144am1	S	R	R	S	S	S	S	S	S	S	S	S	S	924
ODRI	PDR15am1	R	R	S	S	S	S	R	S	R	R	S	S	S	2
ODRI	PDR19am1	S	S	S	S	S	S	R	S	S	S	S	S	S	958
ODRI	PDR1am1	R	R	S	R	S	S	S	S	S	S	S	S	S	89
ODRI	PDR1bm3	R	R	S	R	S	S	S	S	R	S	S	S	S	89
ODRI	PDR21am1	R	R	S	S	S	S	S	S	S	S	S	S	S	88
ODRI	PDR25am1	R	R	S	R	S	S	R	R	S	R	S	S	S	781
ODRI	PDR25am2	R	R	S	R	S	S	R	R	S	R	S	S	S	2
ODRI	PDR25bm1	R	R	S	R	S	S	R	R	S	R	S	S	S	2
ODRI	PDR28am1	R	R	R	R	S	S	R	S	S	R	S	S	R	2
ODRI	PDR33am1	R	R	R	R	S	S	S	R	R	S	R	S	S	2
ODRI	PDR36am1	R	R	R	R	S	S	S	S	S	R	R	S	R	2

ODRI	PDR38bm1	S	S	R	S	S	S	S	S	S	S	S	S	S	57
ODRI	PDR38em1	S	S	R	S	S	S	R	S	S	S	S	S	S	130
ODRI	PDR39am1	R	R	R	R	S	S	R	S	S	R	R	S	S	87
ODRI	PDR6am1	R	R	S	S	S	S	R	S	S	S	S	S	S	88
ODRI	PDR93am1	R	R	S	R	S	S	R	R	S	R	S	S	S	2
ODRI	PDR93bm1	R	R	R	R	S	S	R	R	S	R	R	S	R	2
ODRI	PDR98am1	S	R	R	R	S	S	S	R	R	R	S	S	S	57
ODRI	PDR9am1	R	R	R	S	S	S	R	S	S	S	S	S	S	35
ODRI	TF16m1	R	R	R	S	S	S	R	S	S	S	S	S	S	640
ODRI	TF26m1	R	R	R	S	S	S	R	R	S	R	R	S	R	772
ODRI	TF29m1	S	R	R	S	S	S	S	S	S	S	S	S	S	558
ODRI	TF51m1	R	R	S	R	S	S	R	S	S	S	S	S	S	59
ODRI	TF56m1	S	R	S	S	S	S	S	S	S	S	S	S	S	782
Type Strain	NCTC11047	S	R	S	S	S	S	R	S	R	S	S	S	S	5

Table 5.5. Antimicrobial susceptibility testing and MLST results of colonies of *S. epidermidis* from clinical specimens. S: Susceptible, R: Resistant, I: Intermediate, Cefox: Cefoxitin, Pen: Penicillin, Eryth: Erythromycin, Gent: Gentamicin, Vanc: Vancomycin, Teic: Teicoplanin, Fus: Fusidic Acid, Rif: Rifampicin, Tet: Tetracycline, Cipro: Ciprofloxacin, Clind: Clindamycin, Lzd: Linezolid, Mup: Mupirocin, ODRI: Orthopaedic device-related infection.

5.4.3 Initial Detection of Genetic Determinants of Resistance

5.4.3.1 Detection of acquired resistance genes

Testing for the presence of the 22 resistance genes listed in Table 5.1 in the 87 isolates of *S. epidermidis* (48 listed in Table 5.4 and 39 listed in Table 5.5) generated 1914 observations. Commensal carriage-associated isolates accounted for 1056/1914 (55.2%) of the observations, clinical isolates accounted for 836/1914 (43.7%) and the type strain NCTC 11047 accounted for 22/1914 (1.1%). In concordance with the AST results, there was a lower rate of detection of resistance genes amongst the carriage isolates (14.3%) than the clinical isolates (20.8%). Only two resistance genes were detected in the type strain NCTC 11047 which was isolated from the nose of a healthy individual (4,7). Figure 5.2 shows the prevalence of each resistance gene detected amongst all isolates; genes that were not detected in any isolate are not included in the chart.

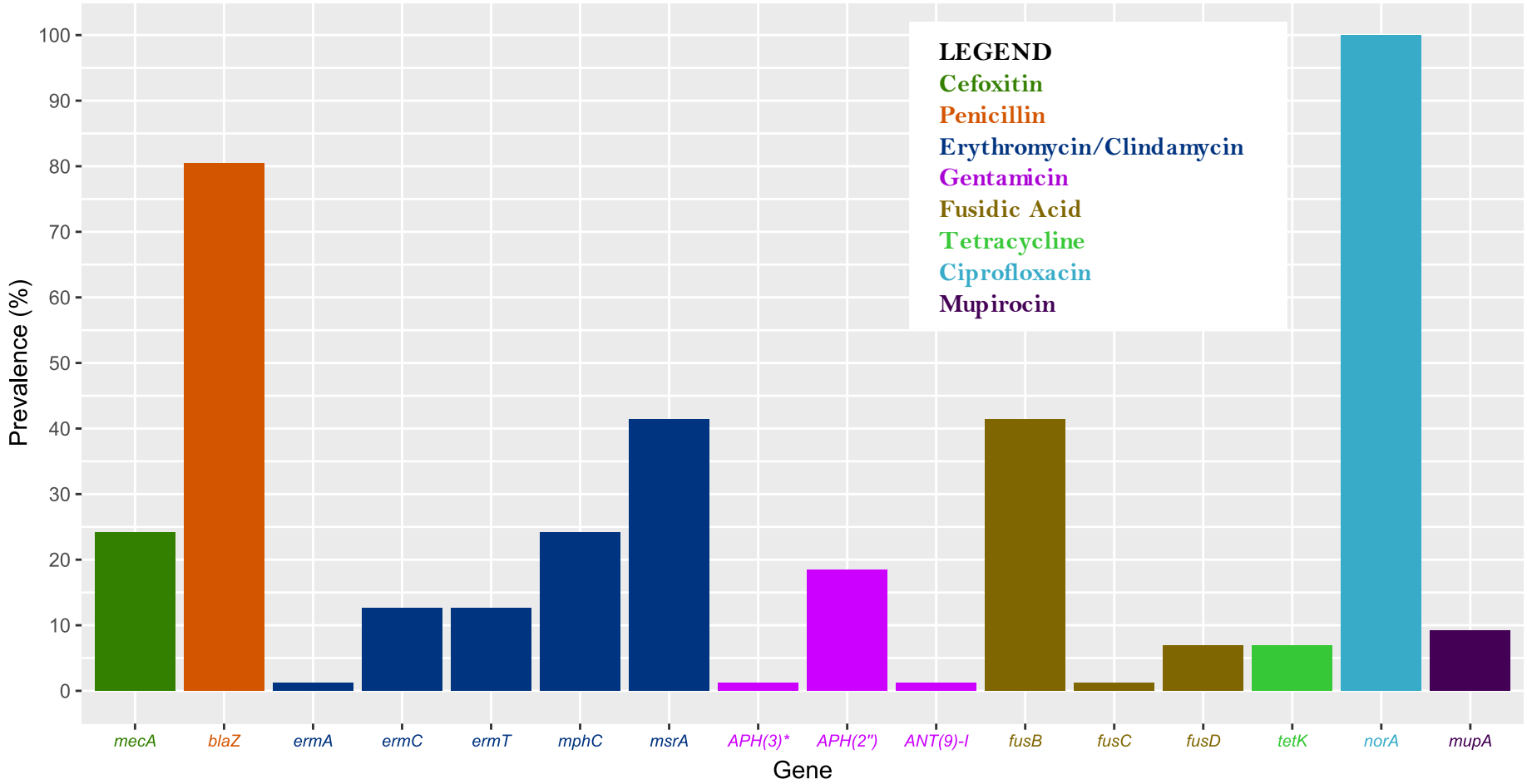


Figure 5.2. Table of prevalence of acquired resistance genes detected by BLAST with a relative coverage >80% in 87 *S. epidermidis* isolates. The barcolour indicates the affected antimicrobial listed in the legend.

5.4.3.2 Detection of amino acid substitutions associated with resistance

The 87 isolates were tested for resistance-conferring amino substitutions affecting ten house-keeping proteins yielding 870 observations. Amongst the commensal carriage isolates 2/480 (0.42%) such substitutions were detected and amongst the clinical isolates 37/380 (9.7%) substitutions were detected. There were no resistance-associated substitutions detected in the type strain NCTC 11047. Figure 5.3 shows prevalence of the amino acid substitution and affected protein for all isolates; substitutions that were not detected were not included in the chart.

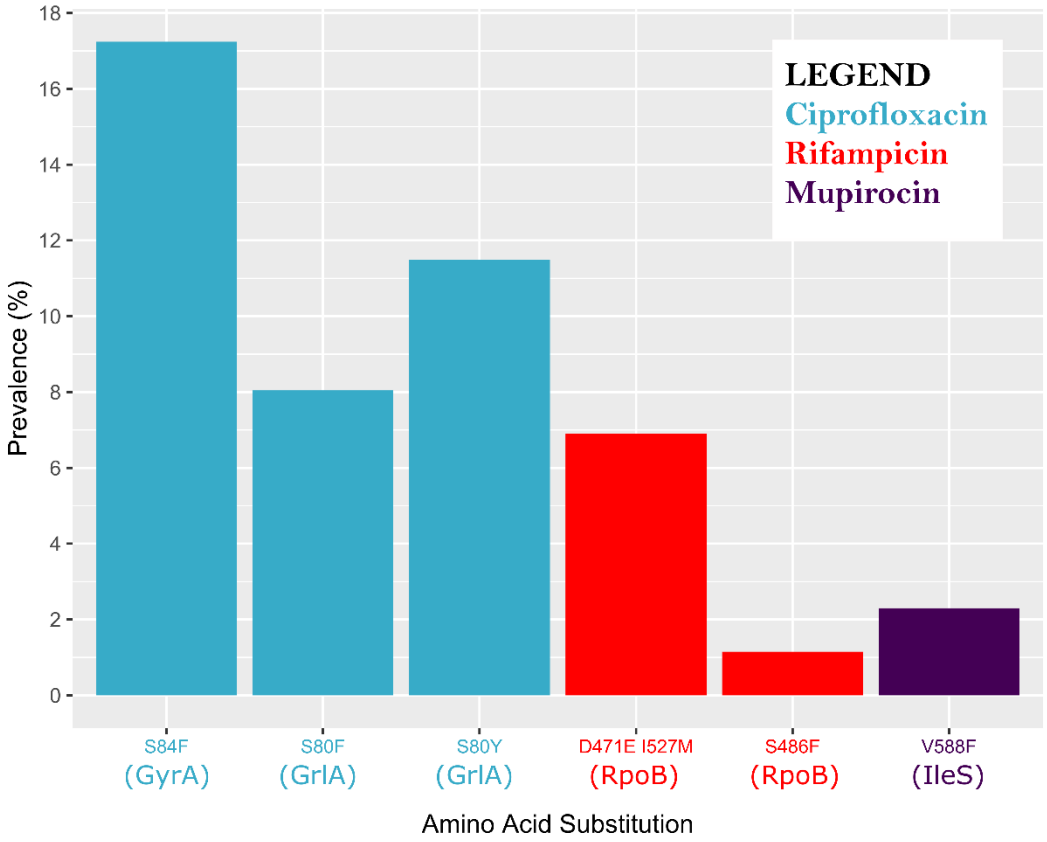


Figure 5.3. Prevalence of amino acid substitutions (with affected protein in parentheses) detected by BLAST in 87 *S. epidermidis* isolates. The bar colour indicates the affected antimicrobial listed in the legend.

5.4.4 Initial Comparison of Phenotyping and Detection of Genetic Determinants of Resistance

To assess the potential ability of a BLAST-based method of predicting antimicrobial resistance in *S. epidermidis* the sensitivity and specificity was determined against the traditional phenotypic methods previously discussed (in Section 5.3.3). In line with previous work taking this approach, a Very Major Error (VME) was defined as the genotyping method not detecting a resistance determinant when there was phenotypic resistance leading to a false-susceptible result and a Major Error (ME) was defined as the detection of a resistance determinant when there was phenotypic susceptibility leading to a false-resistant result (204,376).

Cefoxitin

23/87 (26.4%) of the isolates tested were phenotypically resistant to cefoxitin, amongst which the *mecA* gene was detected in 18/23 (78.3%). Neither *mecA* nor *mecC* were detected in 5/23 (21.7%) of the phenotypically resistant isolates producing a very major error rate of 5.7%. 64/87 (73.6%) of the isolates were phenotypically susceptible to cefoxitin, 2/64 (3.1%) of which had the *mecA* gene resulting in a major error rate of 2.3%.

Penicillin

65/87 (74.7%) of the isolates were phenotypically resistant to penicillin all of which had the resistance-conferring *blaZ* gene. Of the 22 that were phenotypically susceptible, 5/22 (22.7%) harboured the *blaZ* gene with a relative coverage of 94.3% and were therefore regarded as major errors. Two of the phenotypically resistant isolates also had a 94.3% relative coverage of *blaZ* suggesting this was not the cause of discordance.

Erythromycin and Clindamycin

There were 46/87 (52.9%) isolates that were phenotypically resistant to erythromycin, 12/46 (26.1%) of which were also resistant to clindamycin. There was also one isolate that had intermediate resistance to erythromycin and another that had intermediate resistance to clindamycin. Only one isolate harboured the *ermA* gene and thiswas phenotypically resistant to both antimicrobials, this isolate harboured no other genes associated with resistance to the macrolides investigated. The genes *ermC* and *ermT* were always detected together and were present in 10/12 (91.7%) of isolates resistant to both erythromycin and clindamycin. One isolate that had the *ermC* and *ermT* genes was phenotypically susceptible constituting a major error. 36/87 (41.4%) of the isolates harboured the *msrA* gene and 21/36 (58.3%) also had the *mphC* gene. Of these, only one was phenotypically resistant to both macrolides and the rest were resistant to erythromycin only. The isolate that had intermediate resistance to erythromycin by

AST harboured *mphC* and *msrA* and the isolate that had intermediate resistance to clindamycin (and was fully resistant to erythromycin) harboured the *msrA* gene only.

Gentamicin

Phenotypic resistance to gentamicin was observed in 16/87 (18.4%) of the isolates, all of which were of clinical origin and all of which harboured the *aph(2'')* phosphotransferase. One (6.3%) also had the *ant(9)-I* adenylyltransferase and another also had the other phosphotransferase *aph(3')-IIIa*.

Vancomycin and Teicoplanin

Unsurprisingly, none of the isolates were phenotypically or genotypically resistant to the glycopeptides.

Fusidic Acid

35/87 (40.2%) of the isolates were resistant to fusidic acid by traditional AST methods. Of these 34/35 (97.1%) also had the *fusB* gene and 1/35 (2.9%) had the *fusC* gene. There were also three isolates phenotypically susceptible to fusidic acid but had the *fusB* gene and were therefore listed as major errors.

Rifampicin

Phenotypic resistance to rifampicin was observed in 7/87 (8.1%) of the isolates, all of which were from clinical specimens. 6/7 (85.7%) had two amino acid substitutions (D471E and I527M) in the RNA polymerase and one (14.3%) had a single S486F substitution in the RNA polymerase.

Tetracycline

11/87 (12.6%) of the isolates were resistant to tetracycline by AST, of which 5/11 (45.5%) harboured the *tetK* gene. There was also a single isolate which harboured *tetK* but was phenotypically susceptible to tetracycline so was considered a major error.

Ciprofloxacin

All 87 (100%) of the isolates harboured the *norA* gene suggesting this is not a reliable indicator of phenotypic resistance to ciprofloxacin. Gordon *et al.* (2014) observed the same lack of correlation between the presence of *norA* and ciprofloxacin resistance in *S. aureus* (204). 17/87 (19.5%) of the isolates were resistant to ciprofloxacin by phenotype. Of these 8/17 (47.1%) had the S84F amino acid substitution in the GyrA DNA gyrase and S80F in the GrlA DNA topoisomerase, 8/17 (47.1%) had the S84F amino acid substitution in GyrA and S80Y in GrlA, and one (5.8%) had the S80Y in GrlA only.

Linezolid

There were no isolates resistant to linezolid by phenotypic and genotypic detection methods.

Mupirocin

There were 5/87 (5.7%) isolates resistant to mupirocin by disc diffusion phenotype testing, all of which harboured the *mupA* gene. Of the phenotypically susceptible isolates 3/82 (3.6%) harboured the *mupA* gene and two (2.4%) had the V588F amino acid substitution in the Isoleucyl-synthetase (Iles) enzyme and so were defined as major errors.

5.4.4.1 Initial Results

Table 5.6 shows that, in total, there were 1107/1129 (98.1%) concordant results between the genotypic and phenotypic methods for the detection of the presence or absence of antimicrobial resistance. This meant there were 24/1129 (2.1%) discrepancies. Of these, 6/24 (25%) discrepancies were the result of phenotypically resistant isolates lacking any of the resistance-associated determinants tested for. This generated a very major error rate of 0.53%. There were also 18/893 (2.0%) phenotypically susceptible isolates harbouring a resistance-associated determinant. Whilst the overall concordance was high, the number of discrepancies between the methods was high for some of the antibiotics: cefoxitin (7/87 (8.0%)), penicillin (5/87 (5.7%)), mupirocin (5/87 (5.7%)) and moderate for some: fusidic acid (3/87 (3.4%)), clindamycin (2/87 (2.3%)), erythromycin (1/87 (1.1%)) and tetracycline (1.1%)). In addition, there were two phenotypically intermediate results to macrolide antibiotics, one of which was an isolate that harboured two resistance determinants and the other lacked any of the determinants investigated. To address the discrepant or intermediate results additional investigations were undertaken including MIC measurement with E-tests, the biochemical nitrocefin test to determine the presence of a beta-lactamase enzyme (for penicillin) and analysing the sequences of resistance genes for the presence of indels or non-synonymous mutations that could impact translation (383).

Antimicrobial	Resistant Phenotype		Susceptible Phenotype		Intermediate Phenotype		Total	Sensitivity (95% CI)	Specificity (95% CI)	% VME (95% CI)	% ME (95% CI)
	Genotype Present	Genotype Absent	Genotype Present	Genotype Absent	Genotype Present	Genotype Absent					
Cefoxitin	18	5	2	62	0	0	87	0.78 (0.58-0.9)	0.97 (0.89-0.99)	5.7 (2.5-12.8)	2.3 (0.4-8.4)
Penicillin	65	0	5	17	0	0	87	1.0 (0.94-1.0)	0.77 (0.54-0.91)	0.0 (0.0-5.27)	5.75 (2.14-13.51)
Erythromycin	46	0	1	39	1	0	87	1.0 (0.92-1.0)	0.98 (0.85-1.0)	0.0 (0.0-5.27)	1.16 (0.06-7.13)
Gentamicin	16	0	0	71	0	0	87	1.0 (0.81-1.0)	1.0 (0.95-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.27)
Vancomycin	0	0	0	87	0	0	87	N/A	1.0 (0.96-1.0)	N/A	0.0 (0.0-5.27)
Teicoplanin	0	0	0	87	0	0	87	N/A	1.0 (0.96-1.0)	N/A	0.0 (0.0-5.27)
Fusidic Acid	34	0	3	50	0	0	87	1.0 (0.9-1.0)	0.94 (0.83-0.99)	0.0 (0.0-5.27)	3.45 (0.0-5.27)
Rifampicin	7	0	0	80	0	0	87	1.0 (0.56-1.0)	1.0 (0.94-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.27)
Tetracycline	11	0	1	75	0	0	87	1.0 (0.74-1.0)	0.99 (0.91-1.0)	0.0 (0.0-5.27)	1.15 (0.06-7.13)
Ciprofloxacin	17	0	0	70	0	0	87	1.0 (0.82-1.0)	1.0 (0.91-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.28)
Clindamycin	11	1	1	73	0	1	87	0.92 (0.6-0.99)	0.99 (0.91-1.0)	1.15 (0.06-7.13)	1.16 (0.06-7.13)
Linezolid	0	0	0	87	0	0	87	N/A	1.0 (0.91-1.0)	N/A	0.0 (0.0-5.28)
Mupirocin	5	0	5	77	0	0	87	1.0 (0.6-1.0)	0.94 (0.91-0.99)	0.0 (0.0-5.27)	5.75 (2.14-13.51)
Total	**230**	**6**	**18**	**875**	**1**	**1**	**1131**	**0.97 (0.94-0.99)**	**0.98 (0.91-0.99)**	**0.53 (0.24-1.15)**	**1.59 (0.97-2.55)**

Table 5.6. Initial comparison of the presence or absence of a genetic determinant of antimicrobial resistance and the phenotype. VME: Very Major Error, ME: Major Error, N/A: Not Applicable (no resistance observed).

5.4.5 Investigation of Discrepant Results

Further investigations of the discrepant results were performed to try and identify or reconcile the differences between the phenotypic disc diffusion method of AST and the BLAST-based method of detection of genetic determinants of resistance.

Cefoxitin

There were five isolates that were recorded as being resistant to cefoxitin by the disc diffusion method of AST but which lacked the *mecA* or *mecC* genes. Minimum inhibitory concentration testing of oxacillin using E-test strips yielded an MIC of 0.12 mg/L for two of the isolates and 0.25 mg/L for the other three indicating that they were, in fact, susceptible according to BSAC interpretive guidelines (MIC $\leq$2 mg/L) (280).

There were also two isolates which were susceptible to cefoxitin according to disc diffusion testing, but which harboured *mecA*. Testing with oxacillin E-test strips yielded an MIC of 1 mg/L for both isolates, which implies susceptibility according to the BSAC guidelines (280). Further investigation showed that there was 100% relative coverage of the *mecA* gene in both isolate genomes which aligned with the reference gene.

Penicillin

The five isolates that harboured the *blaZ* gene but were phenotypically susceptible to penicillin by disc diffusion were further tested using nitrocefin discs and penicillin E-tests.

Two isolates showed positive nitrocefin hydrolysis and MICs of 0.12 mg/L and 0.25 mg/L confirming that they were, in-fact, phenotypically resistant to penicillin.

One isolate showed no nitrocefin hydrolysis and had an MIC of 0.03 mg/L confirming that it was phenotypically susceptible to penicillin. However, investigation of the *blaZ* gene showed 100% relative coverage using BLAST and no SNVs were detected with ClustalW meaning that the discrepancy could not be resolved using this approach. One of the remaining isolates showed nitrocefin hydrolysis suggesting a functional *blaZ* gene but had a susceptible MIC of 0.015 mg/L according to BSAC guidelines. Moreover, BLASTn yielded relative coverages of 100%, with no SNVs detected by visual analysis of the aligned gene with ClustalW. The other remaining isolate also showed positive nitrocefin hydrolysis and generated a susceptible MIC of 0.06 mg/L. The relative coverage was 94.3% and whilst SNVs were present when aligned with the reference gene there were no indels resulting in frameshifts.

Erythromycin and Clindamycin

One isolate was phenotypically resistant to clindamycin but did not harbour the resistance conferring *ermA*, *ermC*, *ermT* or *lnuA* genes. The isolate did harbour *mphC* and *msrA.* However, whilst these genes confer resistance to erythromycin there is no documented evidence of them conferring resistance to clindamycin, inducible or otherwise. Furthermore, there were no MIC strips available to determine the MIC for clindamycin so this discordance could not be investigated further

An isolate that was susceptible to both erythromycin and clindamycin by disc diffusion, but which harboured the *ermC* gene was further investigated. The *ermC* gene in the test isolate had 99% relative coverage suggesting high homology with the reference gene however when the two genes were aligned with ClustalW there was a five-nucleotide deletion at the 3' end of the coding sequence resulting the loss of an isoleucine residue. A literature search could not confirm if this deletion was the cause of the discrepancy, so this was not resolved.

Fusidic Acid

The three isolates susceptible to fusidic acid but with a positive genotype all had ≥99.8% relative coverage of the *fusB* gene with no non-synonymous mutations, insertions or deletions. There were no E-tests available for MIC testing, so the discordances could not be investigated further.

Tetracycline

One isolate that was susceptible to tetracycline by disc diffusion was also susceptible by E-test method with an MIC of 0.25 mg/L. This isolate also had 100% relative coverage of the *tetK* gene with no non-synonymous mutations, insertions or deletions meaning the results remained discordant.

Mupirocin

Two isolates that were susceptible to mupirocin by the disc diffusion method also had the V588F substitution in the Iles protein. When these isolates underwent E-test MIC testing they both had MICs of 6 mg/L. This concords with low-level mupirocin resistance demonstrated in *S. aureus* that have the same substitution (384).

Three isolates had susceptible phenotypes to mupirocin by disc diffusion and then by E-test (0.125 mg/L) but also had the *mupA* gene. Alignment alongside the reference gene using Clustal W highlighted an adenine nucleotide deletion in position 280 of the three *mupA* genes investigated. This resulted in the same adenine deletion reported by Driscoll *et al.* (2007) which corresponds to position 2441 in the pUSA03 plasmid sequence and causes a frameshift mutation resulting in out-of-frame translation and a premature stop-codon (158).

5.4.5.1 Revised Results

Table 5.7 shows that further investigation resolved several discrepant observations resulting in 1117/1129 (98.9%) concordant results between the phenotypic and genotypic methods. This meant that there were still 12/1129 (1.1%) discrepancies that could not be resolved by MIC testing, biochemical detection of beta-lactamase or inspection of resistance genes. MIC testing improved the sensitivity of the genotypic method of detecting cefoxitin resistance from 78% to 100%. In addition, further analysis improved the specificity of genotypic detection of penicillin resistance and mupirocin resistance from 77% to 85% and 94% to 100% respectively. It was not possible to resolve discrepancies in isolates that were phenotypically susceptible to cefoxitin, erythromycin, fusidic acid, tetracycline and clindamycin but which harboured the genes associated with resistance and therefore there were no changes to the specificities indicated in Table 5.7.

Antimicrobial	Resistant Phenotype		Susceptible Phenotype		Intermediate Phenotype		Total	Sensitivity (95% CI)	Specificity (95% CI)	% VME (95% CI)	% ME (95% CI)
	Genotype Present	Genotype Absent	Genotype Present	Genotype Absent	Genotype Present	Genotype Absent					
Cefoxitin	18	0	2	67	0	0	87	1.0 (0.82-1.0)	0.97 (0.9-0.99)	0.0 (0.0-5.27)	2.3 (0.4-0.84)
Penicillin	67	0	3	17	0	0	87	1.0 (0.95-1.0)	0.85 (0.61-0.96)	0.0 (0.0-5.27)	3.5 (0.9-10.45)
Erythromycin	46	0	1	39	1	0	87	1.0 (0.9-1.0)	0.98 (0.85-1.0)	0.0 (0.0-5.27)	1.16 (0.06-7.13)
Gentamicin	16	0	0	71	0	0	87	1.0 (0.76-1.0)	1.0 (0.95-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.27)
Vancomycin	0	0	0	87	0	0	87	N/A	1.0 (0.96-1.0)	N/A	0.0 (0.0-5.27)
Teicoplanin	0	0	0	87	0	0	87	N/A	1.0 (0.96-1.0)	N/A	0.0 (0.0-5.27)
Fusidic Acid	34	0	3	50	0	0	87	1.0 (0.9-1.0)	0.94 (0.83-0.99)	0.0 (0.0-5.27)	3.45 (0.0-5.27)
Rifampicin	7	0	0	80	0	0	87	1.0 (0.56-1.0)	1.0 (0.94-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.27)
Tetracycline	11	0	1	75	0	0	87	1.0 (0.74-1.0)	0.99 (0.91-1.0)	0.0 (0.0-5.27)	1.15 (0.06-7.13)
Ciprofloxacin	17	0	0	70	0	0	87	1.0 (0.82-1.0)	1.0 (0.91-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.28)
Clindamycin	11	1	1	73	0	1	87	0.92 (0.6-0.99)	0.99 (0.92-1.0)	1.15 (0.06-7.13)	1.16 (0.06-7.13)
Linezolid	0	0	0	87	0	0	87	N/A	1.0 (0.96-1.0)	N/A	0.0 (0.0-5.27)
Mupirocin	7	0	0	77	0	0	87	1.0 (0.56-1.0)	1.0 (0.94-1.0)	0.0 (0.0-5.27)	0.0 (0.0-5.27)
Total	**234**	**1**	**11**	**880**	**1**	**1**	**1131**	1.0 (0.97-1.0)	0.98 (0.97-0.99)	0.09 (0.0-0.57)	0.97 (0.51-1.79)

Figure 5.7. Adjusted comparison of the presence or absence of a genetic determinant of antimicrobial resistance and the phenotype. VME: Very Major Error, ME: Major Error, N/A: Not Applicable (no resistance observed).

5.5 Discussion

The use of whole-genome sequencing to supplement phenotypic methods for the detection and surveillance of antimicrobial resistance in pathogenic organisms has been reported by several researchers (385–387). The implementation of such methods requires rigorous verification against currently adopted methods. Whilst no method may be 100% accurate at predicting antimicrobial resistance *in vivo*, Table 5.8 shows that WGS has been shown to be acceptably sensitive and specific. Verification of the method has already been achieved for some pathogens and WGS has now been adopted as a public health tool (388).

Authors (reference)	Method	Pathogen/s tested (number of isolates)	Concordance with phenotype
Gordon *et al.* (204)	BLAST	*S. aureus* (491)	97% sensitivity, 99% specificity
Pankhurst *et al.* (374)	PhyResSE	*M. tuberculosis* (168)	93% correlation
Eyre *et al.* (375)	BLAST	*N. gonorrhoea* (681)	Same MIC = 53% correlation ± 1 MIC doubling dilution = 93% correlation ± 2 MIC doubling dilution = 98% correlation
Stoesser *et al.* (376)	BLAST	*E. coli* (74)	99% sensitivity, 96% specificity
		K. pneumoniae (69)	95% sensitivity, 97% specificity
Mason *et al.* (291)	BLAST	*S. aureus* (1389)	96.4% sensitivity, 99.2% specificity
	Mykrobe		96.7% sensitivity, 98.9% specificity
	Resfinder		96.7% sensitivity, 99.0% specificity
McDermott *et al.* (389)	BLAST	*Salmonella* (640)	98.8% sensitivity, 99.1% specificity
Zhao *et al.* (390)	BLAST	*Campylobacter coli* (82), *Campylobacter jejuni* (32)	99.2% correlation
Clausen *et al.* (391)	KmerResistance	*Salmonella typhimurium* (48), *E. coli* (48), *E. faecalis* (50), *E. faecium* (50)	96.04% correlation
		E. coli (74), *K. pneumoniae* (69)	99.61% correlation

Table 5.8. Concordance of genotypic and phenotypic methods of determining resistance in common pathogens.

The work outlined in this chapter has shown that for *S. epidermidis*, comparison of the genotypic method against the phenotypic method yielded no discrepant results for gentamicin, rifampicin and ciprofloxacin suggesting the genotypic method is a reliable indicator of the phenotype.

There was one isolate that had intermediate resistance to erythromycin and one that had intermediate resistance to clindamycin by disc diffusion that unfortunately were not MIC tested. An MIC test might have been able to categorise the isolates as either resistant or susceptible or it may have been that they generated an intermediate breakpoint by this method too.

Initial comparison of the two methods yielded six very major errors, five of which were associated with cefoxitin and were resolved by MIC testing with oxacillin. The other very major error was an isolate that was resistant to clindamycin by disc diffusion but lacked *ermA*, *ermC, ermT* or *lnuA*. It is possible that this isolate harboured a different resistant determinant that was not included in the panel (Table 5.1) and therefore not tested however when the isolate genome was interrogated using the online Comprehensive Antibiotic Resistance Database (CARD) (370) no genes associated with clindamycin resistance were detected from the database. Further testing could have included MIC testing, but this was not possible and so the discrepancy remained unresolved.

There were 18 major errors involving seven antibiotics (cefoxitin, penicillin, erythromycin, clindamycin, fusidic acid, tetracycline and mupirocin) in the initial comparison of the two methods. There were three major errors in the 87 isolates tested against fusidic acid in the adjusted comparisons of phenotypic and genotypic methods, however these discordances may have been resolved with MIC testing.

Two of the major errors involving penicillin were resolved by MIC testing, which showed resistance and by performing a nitrocefin test which detected the presence of a beta-lactamase. The other three major errors involving penicillin were not resolved by either phenotypic or genotypic analyses resulting in a major error rate of 3.45% which is higher than the acceptable 3% major error rate specified in the Food and Drug Administration Class II Special Controls Guidance Document: Antimicrobial Susceptibility Test Systems (392). Aubry *et al.* (2020) used the presence of *blaZ* as the standard by which to determine penicillin resistance in *S. epidermidis* and compared this with a zone diameter of less than 26mm, MIC testing using the Vitek2 system (bioMérieux, Marcy l'Etoile, France) and the nitrocefin assay (393). The authors found 4/55 (7.3%) of *blaZ* positive isolates were penicillin susceptible by MIC testing and 5/55 (9.1%) were nitrocefin test negative (393). This work was prompted by the decision of the European Committee on Antimicrobial Susceptibility Testing (EUCAST) to no longer recommend the use

of benzylpenicillin to determine susceptibility in *S. epidermidis* due to poor performance (393,394). Whilst Aubry and colleagues only used PCR to detect the presence of *blaZ* and did not investigate the integrity of the gene the fact that both they and EUCAST have identified inaccuracies demonstrates there is still a gap in the current phenotypic and genotypic methods of determining penicillin resistance.

Two major errors occurred during the comparative testing of cefoxitin which were not resolved by MIC testing or by analysis of the *mecA* gene. The *mecA* gene is regulated by the MecR sensory protein and MecI repressor protein which are encoded by *mecR* and *mecI* respectively (395). Whilst overexpression of *mecI* does not affect methicillin resistance in *S. aureus*, resistance to proteolysis of MecI by MecR2 can cause reduced methicillin resistance (395). It is possible that this phenomenon has occurred in the two *S. epidermidis* isolates resulting in methicillin susceptibility despite the presence of an intact *mecA* gene. Moreover, *blaZ* expression is regulated by a system homologous to that of *mecA* which may have affected the three aforementioned isolates that were phenotypically susceptible to penicillin despite the presence of *blaZ* in their genomes (395). A single isolate that was susceptible to erythromycin and clindamycin by disc diffusion method but harboured the *ermC* gene accounted for the two major errors that involved both antibiotics. Since the *ermC* gene was missing an isoleucine residue and stop codon further investigation into this may help resolve this discrepancy. Two of the five major errors involving mupirocin were resolved by phenotypic determination of low-level resistance by MIC testing alongside identification of the causative genetic element. The other three were resolved by examination of the *mupA* gene which showed a frameshift mutation known to impact expression of MupA in *S. aureus* (158). Interestingly, Driscoll *et al.* (2007) demonstrated that this frameshift was reversible *in vitro* and suggest it could be reversible *in vivo* which could lead to loss of efficacy if using mupirocin to treat or decolonise *S. aureus* (158). Identification of the *mupA* gene and the adenine nucleotide deletion in position 280 may therefore be a more reliable indicator of potential mupirocin resistance than phenotypic methods.

Amongst the 87 isolates tested in this chapter there were none that exhibited phenotypic or genotypic resistance to linezolid and the glycopeptides vancomycin and teicoplanin. This is understandable as resistance to these antibiotics is, whilst emerging, still relatively uncommon in *S. epidermidis* (44,139,396,397). This meant that whilst data on the specificity of the genotypic method (100% concordance with the phenotype) was evaluated there were no means to test the sensitivity.

The results observed in this chapter generally accord with the work of Gordon *et al.* (2014) which had the highest major error rate (5.1%) when predicting penicillin resistance in a validation set of *S. aureus* isolates (204). This was reduced to 0.6% following phenotypic resolution of 22 of the discrepancies by nitrocefin test (n = 11), extended nitrocefin test (n = 3 after two hours) and E-test and nitrocefin test (n = 2) and genotypic analyses of the *blaZ* gene (n = 6) but was still among the highest ME rates in their validation results (204). Overall, Gordon *et al.* (2007) had a higher VME rate at 0.5% and lower ME rate at 0.7% than the results in this thesis chapter. However, their study involved more isolates (n = 491) all of which were of clinical origin.

Further work beyond this thesis should assess performance of genotypic resistance prediction in isolates from a greater range of clinical sources such as blood cultures and venous catheter tips. Testing more clinical isolates would not only add power to the study but would likely yield a greater number of some of the less common resistance phenotypes, e.g. to vancomycin, teicoplanin or linezolid. Furthermore, whilst the phenotypic methods employed in this chapter were intended to emulate those used in the routine clinical practice in the UK, a more comprehensive method of AST would likely yield more accurate results. More comprehensive phenotypic AST could be employed to compare heterogenous resistance as observed with glycopeptides (176,398), and/or intermediate resistance (396) and inducible resistance (20,158). Further genotypic research would likely include the use of more comprehensive databases of resistance-associated genetic elements. This may lead to the detection of resistance elements not previously observed in *S. epidermidis* or other staphylococci. Another interesting field is the characterisation of genetic elements that are associated with resistance to multiple, unrelated antibiotics. Lee and colleagues (2018) have described lineages of *S. epidermidis* that have mutations in the RNA polymerase *rpoB* gene that confer resistance to rifampicin and reduced susceptibility to glycopeptides (182). Further investigation of isolates that harbour resistance determinants but demonstrate phenotypic susceptibility to antimicrobials is also an important area of study. The expression of genetic determinants of resistance may come at a biological cost to an organism and therefore may be regulated (399). Analysis of the bacterial transcriptome to profile resistance gene expression has been performed (400) and maybe a useful tool in describing factors that result in genotypically resistant, phenotypically susceptible isolates. Moreover, animal model or clinical studies determining the relationship between the presence or absence of resistance determinant and drug response may give a more holistic understanding of the treatment of resistant pathogens.

The ability to derive accurate and reliable antimicrobial susceptibility information from whole-genome sequencing data is likely to be a crucial factor if the technology is to be adopted in routine medical microbiology. Successful attempts have already been made involving a range of important pathogens including nontyphoidal *Salmonella*, *Campylobacter spp.*, *S. aureus*, *N. gonorrhoea*, *M. tuberculosis*, *E. coli* and *K. pneumoniae* (374–376,389,390). Since sequencing laboratory methods are now well established, much of the current research is focussed on the bioinformatic methods that can interrogate raw or assembled WGS data to detect genetic determinants of resistance (291,370–372,401,402). Furthermore, the development of curated databases such as CARD, that employs an ontological model to detect any resistance-associated gene, would be useful in previously undocumented resistance determinants in pathogenic species which can occur following horizontal gene transfer. However, since genotypic methods may not detect novel resistance mechanisms it is unlikely that they will fully replace phenotypic methods for the determination of antimicrobial susceptibility.

Here it has been demonstrated that a BLASTn-based method of detecting genetic determinants of resistance in *S. epidermidis* had a 98.76% (95% CI: 97.88 – 99.29%) concordance with the BSAC method of phenotypic AST (280). Table 5.8 shows that this compares favourably with similar work studying other pathogens using BLAST or other methods. Moreover, the overall very major error rate of 0.09% (95CI: 0.0 – 0.57) and major error rate of 0.97% (95% CI: 0.51 – 1.79) (Table 5.7) are within the maximum permissible very major error rate of 1.1% (95% CI: 0.03 – 6.04) and major error rate of 3% required in the Food and Drug Administration Class II Special Controls Guidance Document: Antimicrobial Susceptibility Test Systems (392). The data presented in this chapter are the first demonstration that sequencing based approaches to resistance prediction which have been developed for other species including *Staphylococcus aureus*, can be adapted to reliably predict resistance in *S. epidermidis*. My findings underscore the potential for genotypic assessment to replace phenotypic in routine microbiology practice.

5.6 Conclusion

The work outlined in this chapter has provided the proof-of-principle that interrogation of whole-genome sequence data for the presence of genetic determinants of resistance could be a valid method of antimicrobial susceptibility testing of clinical isolates of *S. epidermidis*. Testing the BLASTn method of detecting resistance determinants yielded a 100% (95% CI: 97 - 100%) sensitivity and 98% (95% CI: 97 – 99%) specificity when compared to BSAC methods. Discrepancies between the BLASTn and BSAC methods were observed for cefoxitin, penicillin, rifampicin, tetracycline, clindamycin and erythromycin that were not resolved and therefore warrant further investigation. Several of the initial discrepancies were resolved by either further phenotypic testing or more in-depth analysis of the genes of interest. If the method were to be validated for clinical, the more in-depth analyses could be programmed into the genotyping to further streamline the process.

Chapter 6
General Discussion and Conclusion

6.1 Discussion

With an aging population and increasing use of implantable medical devices, device infection numbers and incidence rates which are currently are low (0.7 – 4.2%) are set to rise (267). The clinical impact of such infections will increase further as strains of *S. epidermidis* that are resistant to last-line antibiotics are emerging (177,362,403,404) making successful treatment more difficult to achieve. Prevention strategies, including decolonisation regimes, antibiotic prophylaxis, sterile surgical instruments and protocols, and antimicrobial coatings on device surfaces (405), will not completely prevent such infections. Vaccine-based approaches to prevention of staphylococcal have thus far proved beyond reach even for *S. aureus* (406).Thus ODRI diagnosis and management is still an essential requirement in orthopaedic medicine and diagnostic microbiology.

Whole-genome sequencing has already been proven to be a transformative technology in clinical and public health microbiology. The full impact on diagnostics, treatment and infection prevention of WGS is still unknown as researchers pioneer new uses for the technology. WGS has been adopted by Public Health England to provide accredited services for a range of nationally important pathogens (407). These services provide information on the relatedness, virulence and antimicrobial resistance of primary pathogens, where clinical importance has already been determined.

In my research I attempted to determine whether WGS has the potential to discriminate between infecting and non-infecting populations of a type of infection that can be difficult to diagnose. This began with the theory that, since *S. epidermidis* is part of the normal flora of the human skin and infection is the result of accidental contamination of prosthetic devices, then an understanding of the diversity of carriage populations was an important precursor to our understanding of infecting populations.

The first part of my work characterised the diversity of normal human carriage of *S. epidermidis* by measuring the genomic diversity of different strains within and between healthy hosts. Firstly, the commonly used multi-locus sequence typing was applied and detected 32 sequence types, of which seven were novel. Then, genomic relatedness within and between sequence types and participants was determined by three different approaches using pairwise distance measurements:

1. By SNV-based subtyping I determined minimum pairwise distances between genomes from different participants and different sequence types, and demonstrated that multi-locus sequence typing was a stronger indicator of genomic diversity than origin,

2. I found that genomes with a pairwise distance up to 100 SNVs yielded a 95% positive-predictive value that they were isolated from the same participant,
3. Use of maximum-likelihood trees largely yielded results concording with SNV-based subtyping and assigning a positive predictive value, but detected some important outliers where recombination had a more dominant effect than common ancestry and MLST on determining genomic relatedness.

Furthermore, determining fixation indices showed that at a maximum threshold of 550 SNVs (the maximum diversity observed between genomes of the same ST) recombination was seen at a greater degree between isolates from different participants whilst at a minimum threshold of 551 SNVs or no threshold recombination was seen at a greater degree between isolates of different STs.

The second part of my work attempted to develop a genomic-based method of discriminating between infecting and contaminating isolates of *S. epidermidis* cultured from orthopaedic device specimens. A simple method of using mean pairwise distance of genomes from colonies from both same specimen and the same set of specimens was highly discriminatory. The construction of haplotrees based on the biallelic positions of SNVs between colony genomes produced visually demonstrable differences between isolates determined to be associated with device-related infections and isolates determined to be contaminants. An attempt was also made to determine if linear regression could be applied to the increase in diversity (accumulation of SNVs) during the time from orthopaedic device implant to isolation of the infecting *S. epidermidis* colonies. Whilst there was an observable increase in the diversity of infecting *S. epidermidis* colonies over time, a constant correlation could not be determined.

The third part of my work set out to demonstrate that the same genomic data that could potentially aid diagnosis of device-related infection could also be used to guide antibiotic selection. Detection of known genes and mutations associated with antimicrobial resistance concorded with phenotypic antimicrobial susceptibility testing to a high degree of sensitivity and specificity for the antibiotics tested. The 13 antibiotics tested represent some of those which have clinical breakpoints determined by BSAC and EUCAST (280,408) for *S. epidermidis*. The work showed that 12/24 (50%) of the original 24/1129 (2.1%) discrepancies between the genotypic and phenotypic methods were resolved by MIC determination, biochemical detection of the beta-lactamase enzyme or gene interrogation resulting in 1117/1129 (98.9%) concordant results.

6.2 Conclusion

In conclusion, this thesis determines that the while diversity of commensal carriage of *S. epidermidis* is largely driven by a clonality that can be characterised by MLST, analyses using WGS show a more nuanced picture that can be used to measure the relatedness of isolates and infer origin. Similarly, WGS-based measurements of relatedness can be used to accurately discriminate between infecting and contaminating isolates, and visually represent infection. Finally, the same data can be used to reliably detect genomic determinants of antimicrobial resistance.

Chapter 7
References

1. Ogston A. Micrococcus Poisoning. J Anat Physiol. 1882;17(Pt 1):24–58.
2. Rosenbach AJF. Mikro-organismen bei den Wund-infections-krankheiten des Menschen. JF Bergmann; 1884.
3. Welch WH. Conditions underlying the infections of wounds. Am J Med Sci. 1891;102(5):439.
4. Hugh R, Ellis MA. The neotype strain for Staphylococcus epidermidis (Winslow and Winslow 1908) Evans 1916. Int J Syst Evol Microbiol. 1968;18(3):231–9.
5. Fairbrother RW. Coagulase production as a criterion for the classification of the staphylococci. J Pathol Bacteriol. 1940;50(1):83–8.
6. Huebner J, Goldmann DA. Coagulase-negative staphylococci: role as pathogens. Annu Rev Med. 1999;50:223–36.
7. Schleifer KH, Kloos WE. Isolation and characterization of staphylococci from human skin. I. Amended descriptions of Staphylococcus epidermidis and Staphylococcus saprophyticus and descriptions of three new species: Staphylococcus cohnii, Staphylococcus haemolyticus, and. IntJSystBacteriol. 1975 Jan 1;25(1):50–61.
8. Kloos WE, Schleifer KH. Isolation and Characterization of Staphylococci from Human Skin II. Descriptions of Four New Species: Staphylococcus warneri, Staphylococcus capitis, Staphylococcus hominis, and Staphylococcus simulans. Int J Syst Bacteriol. 1975 Jan 1;25(1):62–79.
9. Kloos WE, Schleifer KH. Simplified scheme for routine identification of human Staphylococcus species. J Clin Microbiol. 1975;1(1):82–8.
10. Schoenfelder SMK, Lange C, Eckart M, Hennig S, Kozytska S, Ziebuhr W. Success through diversity - How Staphylococcus epidermidis establishes as a nosocomial pathogen. Int J Med Microbiol. 2010;300(6):380–6.
11. Iwase T, Uehara Y, Shinji H, Tajima A, Seo H, Takada K, et al. Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. Nature. 2010;465(7296):346–9.
12. Pfaller MA, Herwaldt LA. Laboratory, clinical, and epidemiological aspects of coagulase-negative staphylococci. Clin Microbiol Rev. 1988;1(3):281–99.
13. Smith IM, Beals PD, Kingsbury KR, Hasenclever HF. Observations on Staphylococcus albus septicemia in mice and men. AMA Arch Intern Med. 1958;102(3):375–88.
14. Brandt L, Swahn B. Subacute bacterial endocarditis due to coagulase-negative Staphylococcus albus. Acta Med Scand. 1960;166(2):125–32.
15. Freeman J, Platt R, Sidebottom DG, Leclair JM, Epstein MF, Goldmann DA. Coagulase-negative staphylococcal bacteremia in the changing neonatal intensive care unit population: is there an epidemic? Jama. 1987;258(18):2548–52.
16. Rupp ME, Archer GL. Coagulase-negative staphylococci: pathogens associated with

medical progress. Clin Infect Dis. 1994;19(2):231–43.

17. Wisplinghoff H, Bischoff T, Tallent SM, Seifert H, Wenzel RP, Edmond MB. Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. Clin Infect Dis. 2004;39(3):309–17.

18. Jarløv JO, Højbjerg T, Busch-Sørensen C, Scheibel J, Møller JK, Kolmos HJJ, et al. Coagulase-negative Staphylococci in Danish blood cultures: Species distribution and antibiotic susceptibility. J Hosp Infect. 1996;32(3):217–27.

19. Del' Alamo L, Cereda RF, Tosin I, Miranda EA, Sader HS. Antimicrobial susceptibility of coagulase-negative staphylococci and characterization of isolates with reduced susceptibility to glycopeptides. Diagn Microbiol Infect Dis. 1999;34(3):185–91.

20. Gatermann SG, Koschinski T, Friedrich S. Distribution and expression of macrolide resistance genes in coagulase-negative staphylococci. Clin Microbiol Infect. 2007;13(8):777–81.

21. Koksal F, Yasar H, Samasti M. Antibiotic resistance patterns of coagulase-negative staphylococcus strains isolated from blood cultures of septicemic patients in Turkey. Microbiol Res. 2009;164(4):404–10.

22. Jain A, Agarwal A, Verma RK, Awasthi S, Singh KP. Intravenous device associated blood stream staphylococcal infection in paediatric patients. Indian J Med Res. 2011;134(8):193–9.

23. Becker K, Heilmann C, Peters G. Coagulase-negative staphylococci. Clin Microbiol Rev. 2014;27(4):870–926.

24. Ziebuhr W, Hennig S, Eckart M, Kränzler H, Batzilla C, Kozitskaya S. Nosocomial infections by Staphylococcus epidermidis: how a commensal bacterium turns into a pathogen. Int J Antimicrob Agents. 2006;28:14–20.

25. European Center for Disease Prevention and Control. Healthcare-associated infections: surgical site infections. Annual epidemiological report for 2017. ECDC. 2019;(October).

26. WHO. Preventing bloodstream infections from central line venous catheters. URL: http://www. who. int/patientsafety/implementation/bsi/en/index. html. 2012.

27. Shah H, Bosch W, Thompson KM, Hellinger WC. Intravascular catheter-related bloodstream infection. The Neurohospitalist. 2013;3(3):144–51.

28. Han Z, Liang SY, Marschall J. Current strategies for the prevention and management of central line-associated bloodstream infections. Infect Drug Resist. 2010;3:147–63.

29. Otto M. Staphylococcus epidermidis—the'accidental'pathogen. Nat Rev Microbiol. 2009;7(8):555–67.

30. Rogers KL, Fey PD, Rupp ME. Coagulase-Negative Staphylococcal Infections. Infect Dis Clin North Am. 2009;23(1):73–98.

31. Piette A, Verschraegen G. Role of coagulase-negative staphylococci in human disease. Vet Microbiol. 2009;134(1–2):45–54.

32. Klingenberg C, Rønnestad A, Anderson AS, Abrahamsen TG, Zorman J, Villaruz A, et al. Persistent strains of coagulase-negative staphylococci in a neonatal intensive care unit: virulence factors and invasiveness. Clin Microbiol Infect. 2007;13(11):1100–11.

33. Silva GDI De, Kantzanou M, Justice A, Massey RC, Wilkinson AR, Day NPJ, et al. The ica operon and Biofilm Production in Coagulase-Negative Staphylococci Associated with Carriage and Disease in a Neonatal Intensive Care Unit. Microbiology. 2002;40(2):382–

8.

34. Otto C, Cheung GYC, Otto M. Understanding the significance of Staphylococcus epidermidis bacteremia in babies and children. Curr Opin Infect Dis. 2010;23(3):208.

35. Trampuz A, Widmer AF. Infections associated with orthopedic implants. CurrOpinInfectDis. 2006;19:349–56.

36. Pfaller MA, Herwaldt LA. Laboratory, clinical, and epidemiological aspects of coagulase-negative staphylococci. Clin Microbiol Rev. 1988;1(3):281–99.

37. Panda S. Whole-Genome Sequences of Staphylococcus haemolyticus Isolated from Infected Eyes and Healthy Conjunctiva in Bhubaneswar , India. 2016;4(80):1–2.

38. von Eiff C, Peters G, Heilmann C. Review Pathogenesis of infections due to coagulase-negative staphylococci. Lancet Infect Dis. 2002;2:677–85.

39. European Centre for Disease Prevention and Control. Surveillance report- Annual epidemiological report for 2016 - healthcare-associated infections in intensive care units -. ECDC. 2016;(May):11.

40. Nickinson RSJ, Board TN, Gambhir AK, Porter ML, Kay PR. The microbiology of the infected knee arthroplasty. Int Orthop. 2010;34(4):505–10.

41. Flurin L, Greenwood-Quaintance KE, Patel R. Microbiology of polymicrobial prosthetic joint infection. Diagn Microbiol Infect Dis. 2019;94(3):255–9.

42. Lentino JR. Infections Associated with Prosthetic Knee and Prosthetic Hip. Curr Infect Dis Rep. 2004;6(5):388–92.

43. Mohanty SS, Kay PR. Infection in total joint replacements. J Bone Jt Surg. 2004;86(2):266–8.

44. Sanchez A, Benito N, Rivera A, Garcia L, Miro I, Mur I, et al. Pathogenesis of Staphylococcus epidermidis in prosthetic joint infections : can identification of virulence genes differentiate between infecting and commensal strains? J Hosp Infect. 2020;105:561–8.

45. Hellmark B, Söderquist B, Unemo M, Nilsdotter-Augustinsson Å. Comparison of Staphylococcus epidermidis isolated from prosthetic joint infections and commensal isolates in regard to antibiotic susceptibility, agr type, biofilm production, and epidemiology. Int J Med Microbiol. 2013;303(1):32–9.

46. Mohamad M, Deabate L, Belaieff W, Bouvet C, Zingg M, Suva D, et al. Prosthetic Joint Infections Due to Coagulase-Negative Staphylococci. Int J Infect Dis. 2016;3(1):1–8.

47. Hellmark B, Unemo M, Nilsdotter-augustinsson Å, Söderquist B. Antibiotic susceptibility among Staphylococcus epidermidis isolated from prosthetic joint infections with special focus on rifampicin and variability of the rpoB gene. Clin Microbiol Infect. 2009;15(3):238–44.

48. James J, James PJ, Butcher IA, Gardner ER, Hamblen DL, James J. Methicillin-resistant Staphylococcus epidermidis in infection of hip arthroplasties. Bone Joint J. 1994;76(5):725–7.

49. Gill VJ, Selepak ST, Williams EC. Species identification and antibiotic susceptibilities of coagulase-negative staphylococci isolated from clinical specimens. J Clin Microbiol. 1983;18(6):1314–9.

50. Kleeman KT, Bannerman TL, Kloos WE. Species distribution of coagulase-negative staphylococcal isolates at a community hospital and implications for selection of

staphylococcal identification procedures. J Clin Microbiol. 1993;31(5):1318–21.

51. Kawamura Y, Hou XG, Sultana F, Hirose K, Miyake M, Shu SE, et al. Distribution of Staphylococcus species among human clinical specimens and emended description of Staphylococcus caprae. J Clin Microbiol. 1998;36(7):2038–42.

52. Petinaki E, Kontos F, Miriagou V, Maniati M, Hatzi F, Maniatis AN. Survey of methicillin-resistant coagulase-negative staphylococci in the hospitals of central Greece. Int J Antimicrob Agents. 2001;18(6):563–6.

53. Cuevas O, Cercenado E, Vindel A, Guinea J, Sanchez-Conde M, S??nchez-Somolinos M, et al. Evolution of the antimicrobial resistance of Staphylococcus spp. in Spain: Five nationwide prevalence studies, 1986 to 2002. Antimicrob Agents Chemother. 2004;48(11):4240–5.

54. Sivadon V, Rottman M, Chaverot S, Quincampoix JC, Avettand V, De Mazancourt P, et al. Use of genotypic identification by sodA sequencing in a prospective study to examine the distribution of coagulase-negative Staphylococcus species among strains recovered during septic orthopedic surgery and evaluate their significance. J Clin Microbiol. 2005;43(6):2952–4.

55. Mathur T, Singhal S, Khan S, Upadhyay D, Fatma T, Rattan A. Detection of biofilm formation among the clinical isolates of Staphylococci: An evaluation of three different screening methods. Indian J Med Microbiol. 2006;24(1):25.

56. Arciola CR, Campoccia D, An YH, Baldassarri L, Pirini V, Donati ME, et al. Prevalence and antibiotic resistance of 15 minor staphylococcal species colonizing orthopedic implants. Int J Artif Organs. 2006;29(4):395–401.

57. Chaudhury A, Kumar A. In vitro activity of antimicrobial agents against oxacillin resistant staphylococci with special reference to Staphylococcus haemolyticus. Indian J Med Microbiol. 2007;25(1):50.

58. Shin JH, Kim SH, Jeong HS, Oh SH, Kim HR, Lee JN, et al. Identification of Coagulase-Negative Staphylococci Isolated from Continuous Ambulatory Peritoneal Dialysis Fluid using 16S Ribosomal RNA, Tuf and SodA Gene Sequencing. Perit Dial Int. 2011;31(3):340–6.

59. Roach DJ, Burton JN, Lee C, Stackhouse B, Butler-Wu SM, Cookson BT, et al. A Year of Infection in the Intensive Care Unit: Prospective Whole Genome Sequencing of Bacterial Clinical Isolates Reveals Cryptic Transmissions and Novel Microbiota. PLoS Genet. 2015;11(7):1–21.

60. Public Health England. UK Standards for Microbiology Investigations: Investigation of orthopaedic implant associated infections. Bacteriology. 2021;B44(2.1):1–21.

61. Atkins BL, Athanasou N, Deeks JJ, Crook DWM, Simpson H, Timothy EA, et al. Prospective Evaluation of Criteria for Microbiological Diagnosis of Prosthetic-Joint Infection at Revision Arthroplasty Prospective Evaluation of Criteria for Microbiological Diagnosis of Prosthetic-Joint Infection at Revision Arthroplasty. J Clin Microbiol. 1998;36(10):2932–9.

62. Trampuz A, Piper KE, Hanssen AD, Osmon R, Cockerill FR, Steckelberg JM, et al. Sonication of Explanted Prosthetic Components in Bags for Diagnosis of Prosthetic Joint Infection Is Associated with Risk of Contamination Sonication of Explanted Prosthetic Components in Bags for Diagnosis of Prosthetic Joint Infection Is Associated with. Society. 2006;44(2):628–31.

63. Herwaldt LA, Geiss M, Kao C, Pfaller MA. The positive predictive value of isolating

coagulase-negative staphylococci from blood cultures. ClinInfectDis. 1996;22(January):14–20.

64. Souvenir D, Anderson DE, Palpant S, Mroch H, Askin S, Anderson J, et al. Blood cultures positive for coagulase-negative staphylococci: Antisepsis, pseudobacteremia, and therapy of patients. J Clin Microbiol. 1998;36(7):1923–6.

65. Hall KK, Lyman JA. Updated review of blood culture contamination. Clin Microbiol Rev. 2006;19(4):788–802.

66. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization: the true consequences of false-positive results. JAMA. 1991;265(3):365–9.

67. García P, Benítez R, Lam M, Salinas AM, Wirth H, Espinoza C, et al. Coagulase-negative staphylococci: Clinical, microbiological and molecular features to predict true bacteraemia. J Med Microbiol. 2004;53(1):67–72.

68. Tokars JI. Predictive value of blood cultures positive for coagulase-negative staphylococci: implications for patient care and health care quality assurance. Clin Infect Dis. 2004;39(3):333–41.

69. Duck S, Mcdonald LC, Jarvis WR, Mcallister K, Ascp MT, Jerris R, et al. Determining the Significance of Coagulase-Negative Staphylococci Isolated from Blood Cultures at a Community Hospital: A Role for Species and Strain Identification. Infect Control Hosp Epidemiol. 2000;21(03):213–7.

70. Papadimitriou-Olivgeri I, Giormezis N, Papadimitriou-Olivgeris M, Zotou A, Kolonitsiou F, Koutsileou K, et al. Number of positive blood cultures, biofilm formation, and adhesin genes in differentiating true coagulase-negative staphylococci bacteremia from contamination. Eur J Clin Microbiol Infect Dis. 2016;35(1):57–66.

71. Ning Y, Hu R, Yao G, Bo S. Time to positivity of blood culture and its prognostic value in bloodstream infection. Eur J Clin Microbiol Infect Dis. 2016;35(4):619–24.

72. Mirrett S, Weinstein MP, Reimer LG, Wilson ML, Reller LB. Relevance of the number of positive bottles in determining clinical significance of coagulase-negative staphylococci in blood cultures. J Clin Microbiol. 2001;39(9):3279–81.

73. Beekmann SE, Diekema DJ, Doern G V. Determining the clinical significance of coagulase-negative staphylococci isolated from blood cultures. Infect Control Hosp Epidemiol. 2005;26(06):559–66.

74. Healy CM, Baker CJ, Palazzi DL, Campbell JR, Edwards MS. Distinguishing true coagulase-negative Staphylococcus infections from contaminants in the neonatal intensive care unit. J Perinatol. 2012;33(1):1–7.

75. Kassis C, Rangaraj G, Jiang Y, Hachem RY, Raad I. Differentiating culture samples representing coagulase-negative staphylococcal bacteremia from those representing contamination by use of time-to-positivity and quantitative blood culture methods. J Clin Microbiol. 2009;47(10):3255–60.

76. Parvizi J, Tan TL, Goswami K, Higuera C, Della Valle C, Chen AF, et al. The 2018 Definition of Periprosthetic Hip and Knee Infection: An Evidence-Based and Validated Criteria. J Arthroplasty. 2018;33(5):1309–14.

77. Otto M. Staphylococcal Biofilms. Curr Top Microbiol Immunol. 2008;322:207–8.

78. Zhang YQ, Ren SX, Li HL, Wang YX, Fu G, Yang J, et al. Genome-based analysis of virulence genes in a non-biofilm-forming Staphylococcus epidermidis strain (ATCC 12228). Mol Microbiol. 2003;49(6):1577–93.

79. Gill SR, Fouts DE, Archer GL, Mongodin EF, Deboy RT, Ravel J, et al. Insights on Evolution of Virulence and Resistance from the Complete Genome Analysis of an Early Methicillin-Resistant Staphylococcus aureus Strain and a Biofilm-Producing Methicillin-Resistant Staphylococcus epidermidis Strain. J Bacteriol. 2005;187(7):2426–38.

80. Qin L, Da F, Fisher EL, Tan DCS, Nguyen TH, Fu C-L, et al. Toxin Mediates Sepsis Caused by Methicillin-Resistant Staphylococcus epidermidis. PLOS Pathog. 2017;13(2):e1006153.

81. Cheung GYC, Joo H, Chatterjee SS, Otto M. Phenol-soluble modulins – critical determinants of staphylococcal virulence. FEMS Microbiol Rev. 2014;38:698–719.

82. Yao Y, Sturdevant DE, Otto M. Genomewide analysis of gene expression in Staphylococcus epidermidis biofilms: insights into the pathophysiology of S. epidermidis biofilms and the role of phenol-soluble modulins in formation of biofilms. JInfectDis. 2005;191(0022-1899 (Print)):289–98.

83. Le KY, Park MD, Otto M. Immune evasion mechanisms of Staphylococcus epidermidis biofilm infection. Front Microbiol. 2018;9(FEB):1–8.

84. Costerton JWW, Stewart PS, Greenberg EP, Botto M, Fong KY, So AK, et al. Bacterial biofilms: a common cause of persistent infections. Science (80-). 1999;284(5418):1318–22.

85. Roberts AEL, Kragh KN, Bjarnsholt T, Diggle SP, Cope D. The limitations of in vitro experimentation in understanding biofilms and chronic infection. J Mol Biol. 2015;427(23):3646–61.

86. Donlan RM, Costerton JW. Biofilms : Survival Mechanisms of Clinically Relevant Microorganisms. Clin Microbiol. 2002;15(2):167–93.

87. von Eiff C, Peters G, Heilmann C. Review Pathogenesis of infections due to coagulase-negative staphylococci. Lancet Infect Dis. 2002;2(November):677–85.

88. Vuong C, Otto M. Staphylococcus epidermidis infections. Microbes Infect. 2002;4(4):481–9.

89. Fey PD, Olson ME. Current concepts in biofilm formation of Staphylococcus epidermidis. Future Microbiol. 2010;5(6):917–33.

90. Rohde H, Frankenberger S, Zahringer U, Mack D. Structure, function and contribution of polysaccharide intercellular adhesin (PIA) to Staphylococcus epidermidis biofilm formation and pathogenesis of biomaterial-associated infections. Eur J Cell Biol. 2010;89(1):103–11.

91. Vuong C, Voyich JM, Fischer ER, Braughton KR, Whitney AR, DeLeo FR, et al. Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. Cell Microbiol. 2004;6(3):269–75.

92. Singh R, Ray P, Das A, Sharma M. Penetration of antibiotics through Staphylococcus aureus and Staphylococcus epidermidis biofilms. J Antimicrob Chemother. 2010;65(9):1955–8.

93. Rohde H, Burandt EC, Siemssen N, Frommelt L, Burdelski C, Wurster S, et al. Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of Staphylococcus epidermidis and Staphylococcus aureus isolated from prosthetic hip and knee joint infections. Biomaterials. 2007;28(9):1711–20.

94. Rohde H, Burdelski C, Bartscht K, Hussain M, Buck F, Horstkotte MA, et al. Induction of

Staphylococcus epidermidis biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. Mol Microbiol. 2005;55(6):1883–95.

95. Otto M. Virulence Factors of the Coagulase-Negative Staphylococci. Front Biosci. 2004;9:841–63.

96. Rohde H, Kalitzky M, Kröger N, Scherpe S, Horstkotte MA, Knobloch JKM, et al. Detection of virulence-associated genes not useful for discriminating between invasive and commensal Staphylococcus epidermidis strains from a bone marrow transplant unit. J Clin Microbiol. 2004;42(12):5614–9.

97. Galdbart JO, Allignet J, Tung HS, Rydèn C, El Solh N. Screening for Staphylococcus epidermidis markers discriminating between skin-flora strains and those responsible for infections of joint prostheses. J Infect Dis. 2000;182(1):351–5.

98. Frebourg NB, Lefebvre S, Baert S, Lemeland JF. PCR-based assay for discrimination between invasive and contaminating Staphylococcus epidermidis strains. J Clin Microbiol. 2000;38(2):877–80.

99. Conlan S, Mijares LA, Becker J, Blakesley RW, Bouffard GG, Brooks S, et al. Staphylococcus epidermidis pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. Genome Biol. 2012;13(7):R64.

100. Post V, Harris LG, Morgenstern M, Mageiros L, Hitchings MD, Méric G, et al. A comparative genomics study of Staphylococcus epidermidis from orthopedic device-related infections correlated with patient outcome. J Clin Microbiol. 2017;55(10):3089–103.

101. Méric G, Mageiros L, Pensar J, Laabei M, Yahara K, Pascoe B, et al. Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. Nat Commun. 2018;9(1):5034.

102. Pinheiro L, Brito CI et al. Staphylococcus epidermidis and Staphylococcus haemolyticus: detection of biofilm genes and biofilm formation in blood culture isolates from patients in a Brazilian teaching hospital. Diagn Microbiol Infect Dis. 2016;86(1):11–4.

103. Gu J, Li H, Li M, Vuong C, Otto M, Wen Y, et al. Bacterial insertion sequence IS256 as a potential molecular marker to discriminate invasive strains from commensal strains of Staphylococcus epidermidis. J Hosp Infect. 2005;61(4):342–8.

104. Conlon KM, Humphreys H, O'Gara JP. icaR encodes a transcriptional repressor involved in environmental regulation of ica operon expression and biofilm formation in Staphylococcus epidermidis. J Bacteriol. 2002;184(16):4400–8.

105. Rohde H, Kalitzky M, Kroger N, Scherpe S, Horstkotte MA, Knobloch JKM, et al. Detection of virulence-associated genes not useful for discriminating between invasive and commensal Staphylococcus epidennidis strains from a bone marrow transplant unit. J Clin Microbiol. 2004;42(12):5614–9.

106. Giormezis N, Kolonitsiou F, Foka A, Drougka E, Liakopoulos A, Makri A, et al. Coagulase-negative staphylococcal bloodstream and prosthetic-device-associated infections: The role of biofilm formation and distribution of adhesin and toxin genes. J Med Microbiol. 2014;63(2014):1500–8.

107. Götz F. Staphylococcus and biofilms. Mol Microbiol. 2002;43(6):1367–78.

108. Tormo MÁ, Knecht E, Götz F, Lasa I, Penadés JR. Bap-dependent biofilm formation by

pathogenic species of Staphylococcus: Evidence of horizontal gene transfer? Microbiology. 2005;151(7):2465–75.

109. Tolo I, Thomas JC, Fischer RSBB, Brown EL, Gray BM, Robinson DA. Do staphylococcus epidermidis genetic clusters predict isolation sources? J Clin Microbiol. 2016;54(7):1711–9.

110. Harris LG, Murray S, Pascoe B, Bray J, Meric G, Magerios L, et al. Biofilm morphotypes and population Structure among Staphylococcus epidermidis from commensal and clinical samples. PLoS One. 2016;11(3):1–15.

111. Mertens A, Ghebremedhin B. Genetic determinants and biofilm formation of clinical Staphylococcus epidermidis isolates from blood cultures and indwelling devises. Eur J Microbiol Immunol (Bp). 2013;3(2):111–9.

112. Otto M. Staphylococcus aureus and Staphylococcus epidermidis peptide pheromones produced by the accessory gene regulator agr system. Peptides. 2001;22(10):1603–8.

113. Ito T, Okuma K, Ma XX, Yuzawa H, Hiramatsu K. Insights on antibiotic resistance of Staphylococcus aureus from its whole genome: Genomic island SCC. Drug Resist Updat. 2003;6(1):41–52.

114. Zong Z, Peng C, Lü X. Diversity of SCCmec elements in methicillin-resistant coagulase-negative staphylococci clinical isolates. PLoS One. 2011;6(5):1–6.

115. Ito T, Hiramatsu K, Oliveira DC, De Lencastre H, Zhang K, Westh H, et al. Classification of staphylococcal cassette chromosome mec (SCCmec): Guidelines for reporting novel SCCmec elements. Antimicrob Agents Chemother. 2009;53(12):4961–7.

116. Urushibara N, Aung MS, Kawaguchiya M, Kobayashi N. Novel staphylococcal cassette chromosome mec (SCCmec) type XIV (5A) and a truncated SCCmec element in SCC composite islands carrying speG in ST5 MRSA in Japan. J Antimicrob Chemother. 2020 Jan 1;75(1):46–50.

117. Rolo J, de Lencastre H, Miragaia M. Strategies of adaptation of Staphylococcus epidermidis to hospital and community: Amplification and diversification of SCCmec. J Antimicrob Chemother. 2012;67(6):1333–41.

118. Hanssen AM, Ericson Sollid JU. SCCmec in staphylococci: Genes on the move. FEMS Immunol Med Microbiol. 2006;46(1):8–20.

119. Crossley KB, Archer G, Jefferson K, Fowler V. The staphylococci in human disease. New York: Churchill Livingstone; 1997.

120. Cherifi S, Byl B, Deplano A, Nonhoff C, Denis O, Hallin M. Comparative epidemiology of staphylococcus epidermidis isolates from patients with catheter-related bacteremia and from healthy volunteers. J Clin Microbiol. 2013;51(5):1541–7.

121. Barbier F, Lebeaux D, Hernandez D, Delannoy AS, Caro V, François P, et al. High prevalence of the arginine catabolic mobile element in carriage isolates of methicillin-resistant Staphylococcus epidermidis. J Antimicrob Chemother. 2011;66(1):29–36.

122. Miragaia M, de Lencastre H, Perdreau-Remington F, Chambers HF, Higashi J, Sullam PM, et al. Genetic diversity of arginine catabolic mobile element in Staphylococcus epidermidis. PLoS One. 2009;4(11).

123. Shore AC, Rossney AS, Brennan OM, Kinnevey PM, Humphreys H, Sullivan DJ, et al. Characterization of a novel arginine catabolic mobile element (ACME) and staphylococcal chromosomal cassette mec composite island with significant homology to Staphylococcus epidermidis ACME type II in methicillin-resistant Staphylococcus

aureus genotype . Antimicrob Agents Chemother. 2011;55(5):1896–905.

124. O'Connor AM, McManus BA, Coleman DC. First description of novel arginine catabolic mobile elements (ACMEs) types IV and V harboring a kdp operon in Staphylococcus epidermidis characterized by whole genome sequencing. Infect Genet Evol. 2018;61(February):60–6.

125. Granslo HN, Klingenberg C, Fredheim EGA, Rønnestad A, Mollnes TOME, Tromsø N-. Arginine Catabolic Mobile Element Is Associated With Low Antibiotic Resistance and Low Pathogenicity in Staphylococcus epidermidis From Neonates. Pediatr Res. 2010;68(3):237–41.

126. Ferry JG. Formate Dehydrogenase. FEMS Microbiol Rev. 1990;87:1–2.

127. Fournier PE, Gouriet F, Gimenez G, Robert C, Raoult D. Deciphering genomic virulence traits of a staphylococcus epidermidis strain causing native-valve endocarditis. J Clin Microbiol. 2013;51(5):1617–21.

128. Thomas JC, Zhang L, Robinson DA. Differing lifestyles of Staphylococcus epidermidis as revealed through Bayesian clustering of multilocus sequence types. Infect Genet Evol. 2014;22(601):257–64.

129. Espadinha D, Sobral RG, Mendes CI, Méric G, Sheppard SK, Carriço JA, et al. Distinct phenotypic and genomic signatures underlie contrasting pathogenic potential of staphylococcus epidermidis clonal lineages. Front Microbiol. 2019;10(AUG):1–17.

130. Månsson E, Johannesen TB, Nilsdotter-Augustinsson Å, Söderquist B, Stegger M. Comparative genomics of staphylococcus epidermidis from prosthetic-joint infections and nares highlights genetic traits associated with antimicrobial resistance, not virulence. Microb Genomics. 2021;7(2):1–14.

131. Yao Y, Sturdevant DE, Villaruz A, Xu L, Gao Q, Otto M. Factors Characterizing Staphylococcus epidermidis Invasiveness Determined by Comparative Genomics Factors Characterizing Staphylococcus epidermidis Invasiveness Determined by Comparative Genomics. Infect Immun. 2005;73(3):1856–60.

132. Tammelin A, Domicel P, Hambræus A, Ståhle E. Dispersal of methicillin-resistant Staphylococcus epidermidis by staff in an operating suite for thoracic and cardiovascular surgery: Relation to skin carriage and clothing. J Hosp Infect. 2000;44(2):119–26.

133. Gordon RJ, Miragaia M, Weinberg AD, Lee CJ, Rolo J, Giacalone JC, et al. Staphylococcus epidermidis Colonization Is Highly Clonal Across US Cardiac Centers. J Infect Dis. 2012;205:1391–8.

134. Ito T, Hiramatsu K, Tomasz A, De Lencastre H, Perreten V, Holden MTG, et al. Guidelines for reporting novel mecA gene homologues. Antimicrob Agents Chemother. 2012;56(10):4997–9.

135. Harrison EM, Paterson GK, Holden MTG, Larsen J, Stegger M, Larsen AR, et al. Whole genome sequencing identifies zoonotic transmission of MRSA isolates with the novel mecA homologue mecC. EMBO Mol Med. 2013;5(4):509–15.

136. Schwendener S, Cotting K, Perreten V. Novel methicillin resistance gene mecD in clinical Macrococcus caseolyticus strains from bovine and canine sources. Nat Publ Gr. 2017;(November 2016):1–11.

137. Bouchami O, Achour W, Ben Hassen A. Prevalence and mechanisms of macrolide resistance among Staphylococcus epidermidis isolates from neutropenic patients in

Tunisia. Clin Microbiol Infect. 2007;13(1):103–6.

138. Gatermann SG, Koschinski T, Friedrich S. Distribution and expression of macrolide resistance genes in coagulase-negative staphylococci. Clin Microbiol Infect. 2007;13(8):777–81.

139. Deplano A, Vandendriessche S, Nonhoff C, Dodémont M, Roisin S, Denis O. National surveillance of Staphylococcus epidermidis recovered from bloodstream infections in Belgian hospitals. J Antimicrob Chemother. 2016;71(7):1815–9.

140. Lampson BC, Parisi JT. Nucleotide sequence of the constitutive macrolide-lincosamide streptogramin B resistance plasmid pNE131 from Staphylococcus epidermidis and homologies with Staphylococcus aureus plasmids pE194 and pSN2. J Bacteriol. 1986;167(3):888–92.

141. Martineau F, Picard FJ, Lansac N, Ménard C, Roy PH, Ouellette M, et al. Correlation between the Resistance Genotype Determined by Multiplex PCR Assays and the Antibiotic Susceptibility Patterns of Staphylococcus aureus and Staphylococcus epidermidis Correlation between the Resistance Genotype Determined by Multiplex PCR Assay. Antimicrob Agents Chemother. 2000;44(2):231–8.

142. Feßler AT, Wang Y, Wu C, Schwarz S. Plasmid Mobile macrolide resistance genes in staphylococci. Plasmid. 2018;99(April):2–10.

143. Matsuoka M, Inoue M, Endo Y, Nakajima Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol Lett. 2003 Mar 1;220(2):287–93.

144. Benito N, Rivera A, Garcı L, Mur I, Sa A, Gonza Y, et al. Pathogenesis of Staphylococcus epidermidis in prosthetic joint infections: can identification of virulence genes differentiate between infecting and commensal strains? J Hosp Infect. 2020;105(3):561–8.

145. Frey Y, Rodriguez JP, Thomann A, Schwendener S, Perreten V. Genetic characterization of antimicrobial resistance in coagulase-negative staphylococci from bovine mastitis milk. J Dairy Sci. 2013;96(4):2247–57.

146. Castro-Alarcon N, Ribas-Aparicio RM, Silva-Sánchez J, Calderon-Navarro A, Sánchez-Pérez A, Parra-Rojas I, et al. Molecular typing and characterization of macrolide, lincosamide and streptogramin resistance in staphylococcus epidermidis strains isolated in a Mexican hospital. J Med Microbiol. 2011;60(6):730–6.

147. Schmitz F, Fluit AC, Gondolf M, Beyrau R, Lindenlauf E. The prevalence of aminoglycoside resistance and corresponding resistance genes in clinical isolates of staphylococci from 19 European hospitals. J Antimicrob Chemother. 1999;43:253–9.

148. Klingenberg C, Sundsfjord A, Rønnestad A, Mikalsen J, Gaustad P, Flægstad T. Phenotypic and genotypic aminoglycoside resistance in blood culture isolates of coagulase-negative staphylococci from a single neonatal intensive care unit, 1989–2000. 2004;54(5):889–96.

149. Bouchami O, Achour W, Mekni MA, Rolo J, Hassen A Ben. Antibiotic resistance and molecular characterization of clinical isolates of methicillin-resistant coagulase-negative staphylococci isolated from bacteremic patients in oncohematology. Folia Microbiol (Praha). 2011;56(2):122–30.

150. Chopra I, Roberts M. Tetracycline Antibiotics : Mode of Action , Applications , Molecular Biology , and Epidemiology of Bacterial Resistance. Microbiol Mol Biol Rev. 2001;65(2):232–60.

151. Guayt GG, Rothstein DM. Expression of the tetK gene from Staphylococcus aureus in Escherichia coli : Comparison of Substrate Specificities of TetA (B), TetA (C), and TetK Efflux Proteins. Antimicrob Agents Chemother. 1993;37(2):191–8.

152. Jena S, Panda S, Nayak KC, Singh D V. Identification of Major Sequence Types among Multidrug-Resistant Staphylococcus epidermidis Strains Isolated from Infected Eyes and Healthy Conjunctiva. Front Microbiol. 2017;8(August):1–12.

153. Soroush S, Jabalameli F, Taherikalani M, Amirmozafari N, Emaneini M. Microbial Pathogenesis Investigation of biofilm formation ability, antimicrobial resistance and the staphylococcal cassette chromosome mec patterns of methicillin resistant Staphylococcus epidermidis with different sequence types isolated from children. Microb Pathog. 2016;93:126–30.

154. Sanchez-pescador R, Brownt JT, Roberts M, Urdea MS. Homology of the TetM with translational elongation factors: implications for potential modes of tetM conferred tetracycline resistance. Nucleic Acids Res. 1988;16(3):7058.

155. Gilbart J, Perry C, Slocombe B. High-Level Mupirocin Resistance in Staphylococcus aureus : Evidence for Two Distinct Isoleucyl-tRNA Synthetases. Antimicrob Agents Chemother. 1993;37(1):32–8.

156. Ramsey MA, Bradley SF, Kauffman CA, Morton TM. Identification of chromosomal location of mupA gene, encoding low-level mupirocin resistance in staphylococcal isolates. Antimicrob Agents Chemother. 1996;40(12):2820–3.

157. Palepou MFI, Johnson AP, Cookson BD, Beattie H, Charlett A, Woodford N. Evaluation of disc diffusion and Etest for determining the susceptibility of Staphylococcus aureus to mupirocin. J Antimicrob Chemother. 1998;42(5):577–83.

158. Driscoll DG, Young CL, Ochsner UA. Transient loss of high-level mupirocin resistance in Staphylococcus aureus due to MupA polymorphism. Antimicrob Agents Chemother. 2007;51(6):2247–8.

159. Hurdle JG, O'neill AJ, Mody L, Chopra I, Bradley SF. In vivo transfer of high-level mupirocin resistance from Staphylococcus epidermidis to methicillin-resistant Staphylococcus aureus associated with failure of mupirocin prophylaxis. J Antimicrob Chemother. 2005;56(6):1166–8.

160. Liu CT, Hanoian P, French JB, Pringle TH, Hammes-Schiffer S, Benkovic SJ. Functional significance of evolving protein sequence in dihydrofolate reductase from bacteria to humans. Proc Natl Acad Sci. 2013;110(25):10159–64.

161. Schnell JR, Dyson HJ, Wright PE. Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. Annu Rev Biophys Biomol Struct. 2004;33(1):119–40.

162. Darrell JH, Garrod LP, Waterworth PM. Trimethoprim: laboratory and clinical studies. J Clin Pathol. 1968;21(2):202–9.

163. Nurjadi D, Olalekan AO, Layer F, Shittu AO, Alabi A, Ghebremedhin B, et al. Emergence of trimethoprim resistance gene dfrG in Staphylococcus aureus causing human infection and colonization in sub-Saharan Africa and its import to Europe. J Antimicrob Chemother. 2014;69(9):2361–8.

164. Kadlec K, Schwarz S. Identification of a novel trimethoprim resistance gene, dfrK, in a methicillin-resistant Staphylococcus aureus ST398 strain and its physical linkage to the tetracycline resistance gene tet(L). Antimicrob Agents Chemother. 2009;53(2):776–8.

165. Totake K, Kobayashi N, April A. Trimethoprim Resistance and in Staphylococcus

epidermidis Susceptibility Genes. Microbiol Immunol. 1998;42(7):497–501.

166. Fowler PW, Cole K, Gordon NC, Kearns AM, Llewelyn MJ, Peto TEA, et al. Robust Prediction of Resistance to Trimethoprim in Staphylococcus aureus. Cell Chem Biol. 2018;25:1–11.

167. Dale GE, Broger C, Arcy AD, Hartman PG, Dehoogt R, Jolidon Á, et al. A Single Amino Acid Substitution in Staphylococcus aureus Dihydrofolate Reductase Determines Trimethoprim Resistance. J Mol Biol. 1997;266:23–30.

168. Ruppé E, Lazarevic V, Girard M, Mouton W, Ferry T, Laurent F, et al. Clinical metagenomics of bone and joint infections: a proof of concept study. Nat Sci Reports. 2017;7(1):7718.

169. Nagarajan R. MiniReview: Antibacterial Activities and Modes of Action of Vancomycin and Related Glycopeptides. Antimicrob Agents Chemother. 1991;35(4):605–9.

170. Srinivasan A, Dick JD, Perl TM, Tenover FC, Biddle JW, Lancaster M V., et al. Vancomycin Resistance in Staphylococci. Clin Microbiol Rev. 2002;15(3):430–8.

171. Biavasco F, Vignaroli C, Varaldo PE. Glycopeptide Resistance in Coagulase-Negative Staphylococci. Eur J Clin Microbiol Infect Dis. 2000;19:403–17.

172. Cremniter J, Slassi A, Quincampoix JC, Sivadon-Tardy V, Bauer T, Porcher R, et al. Decreased susceptibility to teicoplanin and vancomycin in coagulase-negative Staphylococci isolated from orthopedic-device-associated infections. J Clin Microbiol. 2010;48(4):1428–31.

173. Gazzola S, Cocconcelli PS. Vancomycin heteroresistance and biofilm formation in Staphylococcus epidermidis from food. Microbiology. 2008;154(10):3224–31.

174. Watanabe Y, Cui L, Katayama Y, Kozue K, Hiramatsu K. Impact of rpoB Mutations on Reduced Vancomycin Susceptibility in Staphylococcus aureus. J Clin Microbiol. 2011;49(7):2680–4.

175. Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, et al. Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing. ProcNatlAcadSciUSA. 2007;104(22):9451–6.

176. Paula A, Nunes F, Iorio P, Calleg C, Bastos R, Sousa L De. Heterogeneous resistance to vancomycin in Staphylococcus epidermidis , Staphylococcus haemolyticus and Staphylococcus warneri clinical strains : characterisation of glycopeptide susceptibility profiles and cell wall thickening. Int J Antimicrob Agents. 2006;27:307–15.

177. Cremniter J, Sivadon-tardy V, Caulliez C, Bauer T, Porcher R, Lortat-jacob A. Genetic Analysis of Glycopeptide-Resistant Staphylococcus epidermidis Strains from Bone and Joint Infections. J Clin Microbiol. 2013;51(3):1014–9.

178. McCann MT, Gilmore BF, Gorman SP. Staphylococcus epidermidis device-related infections : pathogenesis and clinical management. J Pharm Pharmacol. 2008;60:1551–71.

179. Boyle DL, Takemoto L, Brady JP, Wawrousek EF. Characterization of coagulase negative staphylococcal isolates from blood with reduced susceptibility to glycopeptides and therapeutic options. BMC Infect Dis. 2009;9(83):1–10.

180. Wehrli W. Rifampin: mechanisms of action and resistance. Rev Infect Dis. 1983;5(Supplement_3):S407–11.

181. Mi Y, Greenwood-quaintance KE, Brinkman CL, Lee JYH, Howden BP, Patel R. Rifampicin

resistance in Staphylococcus epidermidis : molecular characterisation and fitness cost of rpoB mutations. Int J Antimicrob Agents. 2018;51:670–7.

182. Lee JYH, Monk IR, da Silva AG, Seemann T, Chua KYL, Kearns A, et al. Global spread of three multidrug-resistant lineages of Staphylococcus epidermidis Supplementary Material. Nat Microbiol. 2018;3(10):1175.

183. Haunreiter VD, Boumasmoud M, Häffner N, Wip D, Leimer N, Rachmühl C, et al. In-host evolution of Staphylococcus epidermidis in a pacemaker-associated endocarditis resulting in increased antibiotic tolerance. Nat Commun. 2019;10(1):1149.

184. Szczuka E, Kaznowski A. Antimicrobial activity of tigecycline alone or in combination with rifampin against Staphylococcus epidermidis in biofilm. Folia Microbiol (Praha). 2014;59:283–8.

185. Zheng Z, Stewart PS. Penetration of Rifampin through Staphylococcus epidermidis Biofilms. Antimicrob Agents Chemother. 2002;46(3):900–3.

186. Dunne WM, Mason E, Kaplan SL. Diffusion of Rifampin and Vancomycin through a Staphylococcus epidermidis Biofilm. Antimicrob Agents Chemother. 1993;37(12):2522–6.

187. Zheng Z, Stewart PS. Growth limitation of Staphylococcus epidermidis in biofilms contributes to rifampin tolerance. Biofilms. 2004;1:31–5.

188. Chaieb K, Zmantar T, Chehab O, Boucham O, Hasen A Ben, Mahdouani K, et al. Antibiotic resistance genes detected by multiplex PCR assays in Staphylococcus epidermidis strains isolated from dialysis fluid and needles in a dialysis service. Jpn J Infect Dis. 2007;60(4):183–7.

189. Bender J, Strommenger B, Steglich M, Zimmermann O, Fenner I, Lensing C, et al. Linezolid resistance in clinical isolates of Staphylococcus epidermidis from German hospitals and characterization of two cfr-carrying plasmids. J Antimicrob Chemother. 2015;70(6):1630–8.

190. Bozdogan B, Appelbaum PC. Oxazolidinones : activity , mode of action , and mechanism of resistance. Int J Antimicrob Agents. 2004;23:113–9.

191. Livermore DM. Linezolid in vitro : mechanism and antibacterial spectrum. J Antimicrob Chemother. 2003;51:9–16.

192. Thompson J, O'Connor M, Mills JA, Dahlberg AE. The protein synthesis inhibitors, oxazolidinones and chloramphenicol, cause extensive translational inaccuracy in vivo. J Mol Biol. 2002/09/10. 2002;322(2):273–9.

193. Hong T, Li X, Wang J, Sloan C, Cicogna C. Sequential linezolid-resistant Staphylococcus epidermidis isolates with G2576T mutation. J Clin Microbiol. 2007;45(10):3277–80.

194. Long KS, Vester B. Resistance to Linezolid Caused by Modifications at Its Binding Site on the Ribosome. Antimicrob Agents Chemother. 2012;603–12.

195. Mendes RE, Deshpande LM, Costello AJ, Farrell DJ. Molecular epidemiology of Staphylococcus epidermidis clinical isolates from U.S. hospitals. Antimicrob Agents Chemother. 2012;56(9):4656–61.

196. May L, Klein EY, Rothman RE, Laxminarayan R. Trends in antibiotic resistance in coagulase-negative staphylococci in the United States, 1999 to 2012. Antimicrob Agents Chemother. 2014;58(3):1404–9.

197. Baos E, Candel FJ, Merino P, Pena I, Picazo JJ. Characterization and monitoring of

linezolid-resistant clinical isolates of Staphylococcus epidermidis in an intensive care unit 4 years after an outbreak of infection by cfr-mediated linezolid-resistant Staphylococcus aureus. Diagn Microbiol Infect Dis. 2013;76(3):325–9.

198. Bonilla H, Huband MD, Seidel J, Schmidt H, Lescoe M, McCurdy SP, et al. Multicity Outbreak of Linezolid-Resistant *Staphylococcus epidermidis* Associated with Clonal Spread of a *cfr* -Containing Strain. Clin Infect Dis. 2010;51(October):796–900.

199. Hooper DC. Mechanisms of Action of Antimicrobials : Focus on Fluoroquinolones. Clin Infect Dis. 2001;32(Suppl 1):9–15.

200. Hooper DC. Mode of Action of Fluoroquinolones. Drugs. 1999;58:6–10.

201. Hooper DC. Emerging Mechanisms of Fluoroquinolone Resistance. Emerg Infect Dis. 2001;7(2):337–41.

202. Yamada M, Yoshida J, Hatou S, Yoshida T, Minagawa Y. Mutations in the quinolone resistance determining region in Staphylococcus epidermidis recovered from conjunctiva and their association with susceptibility to various fluoroquinolones. Br J Ophthalmol. 2008;92(6):848–51.

203. Wang T, Tanaka M, Product N, Co DP. Detection of grlA and gyrA Mutations in 344 Staphylococcus aureus Strains. Antimicrob Agents Chemother. 1998;42(2):236–40.

204. Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, et al. Prediction of staphylococcus aureus antimicrobial resistance by whole-genome sequencing. J Clin Microbiol. 2014;52(4):1182–91.

205. Li Z, Deguchi T, Yasuda M, Kawamura T, Kanematsu E, Nishino Y, et al. Alteration in the GyrA Subunit of DNA Gyrase and the ParC Subunit of DNA Topoisomerase IV in Quinolone-Resistant Clinical Isolates of Staphylococcus epidermidis. Antimicrob Agents Chemother. 1998;42(12):3293–5.

206. Sreedharan S, Peterson LR, Fisher LM. Ciprofloxacin resistance in coagulase-positive and -negative staphylococci: Role of mutations at serine 84 in the DNA gyrase A protein of Staphylococcus aureus and Staphylococcus epidermidis. Antimicrob Agents Chemother. 1991;35(10):2151–4.

207. Otto J, Høiby N, Jarløv JO, Kemp M, Tvede M, Bangsborg JM, et al. Excretion of ciprofloxacin in sweat and multiresistant Staphylococcus epidermidis. Lancet. 1997;349(9046):167–9.

208. Castanheira M, Watters AA, Bell JM, Turnidge JD, Jones RN. Fusidic Acid Resistance Rates and Prevalence of Resistance Mechanisms among Staphylococcus spp . Isolated in North. Antimicrob Agents Chemother. 2010;54(9):3614–7.

209. Farrell DJ, Castanheira M, Chopra I. Characterization of Global Patterns and the Genetics of Fusidic Acid Resistance. Clin Infect Dis. 2011;52(Suppl 7):487–92.

210. Mclaws F, Chopra I, O'Neill AJ. High prevalence of resistance to fusidic acid in clinical isolates of Staphylococcus epidermidis. J Antimicrob Chemother. 2008;61(5):1040–3.

211. Besier S, Ludwig A, Brade V, Wichelhaus TA. Molecular analysis of fusidic acid resistance in Staphylococcus aureus. Mol Microbiol. 2003;47:463–9.

212. Hung WC, Chen HJ, Lin YT, Tsai JC, Chen CW, Lu HH, et al. Skin commensal staphylococci may act as reservoir for fusidic acid resistance genes. PLoS One. 2015;10(11):1–15.

213. Norstrom T, Lannergård J, Hughes D. Genetic and Phenotypic Identification of Fusidic

Acid-Resistant Mutants with the Small-Colony-Variant Phenotype in Staphylococcus aureus . Antimicrob Agents Chemother. 2007;51(12):4438–46.

214. Turnidge J, Collignon P. Resistance to fusidic acid. Int J Antimicrob Agents. 1999;12:35–44.

215. Carbonnelle E, Beretti JL, Cottyn S, Quesne G, Berche P, Nassif X, et al. Rapid identification of staphylococci isolated in clinical microbiology laboratories by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol. 2007;45(7):2156–61.

216. Dubois D, Leyssene D, Chacornac JP, Kostrzewa M, Schmit PO, Talon R, et al. Identification of a variety of Staphylococcus species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol. 2010;48(3):941–5.

217. Nagel JL, Huang AM, Kunapuli A, Gandhi TN, Washer LL, Lassiter J, et al. Impact of antimicrobial stewardship intervention on coagulase-negative Staphylococcus blood cultures in conjunction with rapid diagnostic testing. J Clin Microbiol. 2014;52(8):2849–54.

218. Harris LG, El-Bouri K, Johnston S, Rees E, Frommelt L, Siemssen N, et al. Rapid identification of staphylococci from prosthetic joint infections using MALDI-TOF mass-spectrometry. Int J Artif Organs. 2010;33(9):568–74.

219. De Buyser M-LL, Morvan A, Aubert S, Dilasser F, el Solh N. Evaluation of a ribosomal RNA gene probe for the identification of species and subspecies within the genus Staphylococcus. J Gen Microbiol. 1992;138(5):889–99.

220. Takahashi T, Satoh I, Kikuchi N. Phylogenetic relationships of 38 taxa of the genus Staphylococcus based on 16S rRNA gene sequence analysis. Int J Syst Bacteriol. 1999;49(1999):725–8.

221. Ghebremedhin B, Layer F, König W, König B. Genetic classification and distinguishing of Staphylococcus species based on different partial gap, 16S rRNA, hsp60, rpoB, sodA, and tuf gene sequences. J Clin Microbiol. 2008;46(3):1019–25.

222. Kosecka-Strojek M, Sabat AJ, Akkerboom V, Becker K, van Zanten E, Wisselink G, et al. Development and Validation of a Reference Data Set for Assigning Staphylococcus Species Based on Next-Generation Sequencing of the 16S-23S rRNA Region. Front Cell Infect Microbiol. 2019;9(278):1–19.

223. Shin JH, Kim SH, Jeong HS, Oh SH, Kim HR, Lee JN, et al. Identification of coagulase-negative staphylococci isolated from continuous ambulatory peritoneal dialysis fluid using 16S ribosomal RNA, tuf, and SodA gene sequencing. Perit Dial Int. 2011;31(3):340–6.

224. Heikens E, Fleer a, Paauw a, Florijn a, Fluit a C. Comparison of Genotypic and Phenotypic Methods for Species-Level Identification of Clinical Isolates of Coagulase-Negative Staphylococci Comparison of Genotypic and Phenotypic Methods for Species-Level Identification of Clinical Isolates of Coagulase-Nega. J Clin Microbiol. 2005;43(5):2286–90.

225. Poyart C, Quesne G, Boumaila C, Trieu-Cuot P. Rapid and Accurate Species-Level Identification of Coagulase- Negative Staphylococci by Using the. J Clin Microbiol. 2001;39(12):4296–301.

226. Mellmann A, Becker K, Von Eiff C, Keckevoet U, Schumann P, Harmsen D. Sequencing and staphylococci identification. Emerg Infect Dis. 2006;12(2):333–6.

227. Giammarinaro P, Leroy S, Chacornac JP, Delmas J, Talon R. Development of a new oligonucleotide array to identify staphylococcal strains at species level. J Clin Microbiol. 2005;43(8):3673–80.

228. Wang H, Du P, Li J, Zhang Y, Zhang W, Han N, et al. Comparative analysis of microbiome between accurately identified 16S rDNA and quantified bacteria in simulated samples. J Med Microbiol. 2014;63(3):433–40.

229. Baird-Parker AC. A classification of micrococci and staphylococci based on physiological and biochemical tests. J Gen Microbiol. 1963;30:409–27.

230. Götz F, Bannerman T, Schleifer K-H. The genera staphylococcus and macrococcus. In: The prokaryotes. Springer; 2006. p. 5–75.

231. Struelens MJ, Study E, Markers E, Microbiology C, Escmid ID. Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems. Clin Microbiol Infect. 1996;2(1):2–11.

232. van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of Genomic Typing in Taxonomy, Evolutionary Genetics, and Microbial Epidemiology. Clin Microbiol Rev. 2001 Jul 1;14(3):547–60.

233. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A. 1998;95(6):3140–5.

234. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST : Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data. 2004;186(5):1518–30.

235. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 2018 Sep 24;3:124.

236. Miragaia M, Carrico JA, Thomas JC, Couto I, Enright MC, De Lencastre H, et al. Comparison of molecular typing methods for characterization of Staphylococcus epidermidis: proposal for clone definition. J Clin Microbiol. 2008;46(1):118–29.

237. Ichiyama S, Ohta M, Shimokata K, Kato N, Takeuchi J. Genomic DNA fingerprinting by pulsed-field gel electrophoresis as an epidemiological marker for study of nosocomial infections caused by methicillin-resistant Staphylococcus aureus. J Clin Microbiol. 1991;29(12):2690–5.

238. Miragaia M, Thomas JC, Couto I, Enright MC, De Lencastre H. Inferring a population structure for Staphylococcus epidermidis from multilocus sequence typing data. J Bacteriol. 2007;189(6):2540–52.

239. Sloos JH, Dijkshoorn L, Vogel L, Van Boven CPA. Performance of phenotypic and genotypic methods to determine the clinical relevance of serial blood isolates of Staphylococcus epidermidis in patients with septicemia. J Clin Microbiol. 2000;38(7):2488–93.

240. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol. 2015;53(4):1072–9.

241. Galdbart JO, Morvan A, Desplaces N, El Solh N. Phenotypic and genomic variation among Staphylococcus epidermidis strains infecting joint prostheses. J Clin Microbiol. 1999;37(5):1306–12.

242. Rijnders BJA, Van Wijngaerden E, Van Eldere J, Peetermans WE. Polyclonal Staphylococcus epidermidis intravascular catheter-related infections. Clin Microbiol Infect. 2001;7(7):388–91.

243. Mickelsen PA, Plorde JJ, Gordon KP, Hargiss C, McClure J, Schoenknecht FD, et al. Instability of antibiotic resistance in a strain of Staphylococcus epidermidis isolated from an outbreak of prosthetic valve endocarditis. JInfectDis. 1985;152(1):50–8.

244. Ziebuhr W, Dietrich K, Trautmann M, Wilhelm M. Chromosomal rearrangements affecting biofilm production and antibiotic resistance in a *Staphylococcus epidermidis* strain causing shunt-associated ventriculitis. Int J Med Microbiol. 2000;290(1):115–20.

245. Sharma M, Riederer K, Johnson LB, Khatib R. Molecular analysis of coagulase-negative Staphylococcus isolates from blood cultures: prevalence of genotypic variation and polyclonal bacteremia. Clin Infect Dis. 2001;33(8):1317–23.

246. Thomas JC, Vargas MR, Miragaia M, Peacock SJ, Archer GL, Enright MC. Improved multilocus sequence typing scheme for Staphylococcus epidermidis. J Clin Microbiol. 2007;45(2):616–9.

247. arcC - Carbamate kinase - Staphylococcus epidermidis - arcC gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/A0A223LC81

248. aroE - Shikimate dehydrogenase (NADP(+)) - Staphylococcus epidermidis (strain ATCC 35984 / RP62A) - aroE gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/Q5HNV1

249. gtr - ABC transporter - Staphylococcus epidermidis - gtr gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/J7F8M2

250. mutS - DNA mismatch repair protein MutS - Staphylococcus epidermidis - mutS gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/A0A4V2KVF2

251. pyrR - Bifunctional protein PyrR - Staphylococcus epidermidis - pyrR gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/A0A2G7I217

252. tpiA - Triosephosphate isomerase - Staphylococcus epidermidis - tpiA gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/A0A4Q9W4W7

253. yqiL - Acetyl coenzyme A acetyltransferase - Staphylococcus epidermidis - yqiL gene & protein [Internet]. [cited 2021 Jun 21]. Available from: https://www.uniprot.org/uniprot/Q6YBD3

254. Allele Sequences [Internet]. [cited 2021 Jun 21]. Available from: https://pubmlst.org/bigsdb?db=pubmlst_sepidermidis_seqdef&page=downloadAlleles

255. Kozitskaya S, Olson ME, Fey PD, Witte W, Ohlsen K, Ziebuhr W, et al. Clonal Analysis of Staphylococcus epidermidis Isolates Carrying or Lacking Biofilm-Mediating Genes by Multilocus Sequence Typing. J Clin Microbiol. 2005;43(9):4751–7.

256. Månsson E, Hellmark B, Sundqvist M, Söderquist B. Sequence types of Staphylococcus epidermidis associated with prosthetic joint infections are not present in the laminar airflow during prosthetic joint surgery. Apmis. 2015;123(7):589–95.

257. Both A, Huang J, Qi M, Lausmann C, Weißelberg S, Büttner H, et al. Distinct clonal lineages and within-host diversification shape invasive staphylococcus epidermidis populations. Vol. 17, PLoS Pathogens. 2021. 1–39 p.

258. Goering RV, Winters MA. Rapid Method for Epidemiological Evaluation of Gram-Positive Cocci by Field Inversion Gel Electrophoresis. J Clin Microbiol. 1992;30(3):577–80.

259. Lina B, Vandenesch F, Etienne J, Kreiswirth B, Fleurette J. Comparison of coagulase-negative staphylococci by pulsed-field gel electrophoresis. FEMS Microbiol Lett. 1992;92(2):133–8.

260. Oh J, Byrd AL, Deming C, Conlan S, Program NCS, Kong HH, et al. Biogeography and individuality shape function in the human skin metagenome. Nature. 2014;514(7520):59.

261. Oh J, Byrd AL, Park M, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell. 2016;165(4):854–66.

262. Su F, Tian R, Yang Y, Li H, Sun G, Li Y, et al. Comparative Genome Analysis Reveals the Molecular Basis of Niche Adaptation of Staphylococcus epidermidis Strains. Front Genet. 2020;11(November):1425.

263. Tapscott DC WC. Orthopedic Implant Materials. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020.

264. Phillips JE, Crane TP, Noy M, Elliott TSJ, Grimer RJ. The incidence of deep prosthetic infections in specialist orthopaedic hospital - A 15-year prospective survey. J Bone Jt Surgery-British Vol. 2006;88(7):943–8.

265. Steinmetz S, Wernly D, Moerenhout K, Trampuz A, Borens O. Infection after fracture fixation. EFORT Open Rev. 2019;4(7):468–75.

266. Zimmerli W, Trampuz A, Ochsner PE. Prosthetic-Joint Infections. N Engl J Med. 2004;351(16):1645–54.

267. Moriarty TF, Kuehl R, Coenye T, Metsemakers WJ, Morgenstern M, Schwarz EM, et al. Orthopaedic device-related infection: Current and future interventions for improved prevention and treatment. EFORT Open Rev. 2016;1(4):89–99.

268. Costerton JW, Stewart PS, Greenberg EP. Bacterial Biofilms: A Common Cause of Persistent Infections. Microbes, Immun Dis. 1999;284(314):1318–22.

269. Yao Y, Sturdevant DE, Otto M. Genomewide analysis of gene expression in _Staphylococcus epidermidis_ biofilms: insights into the pathophysiology of _S. epidermidis_ biofilms and the role of phenol-soluble modulins in formation of biofilms. JInfectDis. 2005;191(0022-1899 (Print)):289–98.

270. Qasim SN, Swann A, Ashford R. The DAIR (debridement, antibiotics and implant retention) procedure for infected total knee replacement – a literature review. Sicot-J. 2017;3:2.

271. Cooper HJ, Della Valle CJ. The two-stage standard in revision total hip replacement. Bone Joint J. 2013;95 B(11 Suppl A):84–7.

272. Metsemakers WJ, Morgenstern M, McNally MA, Moriarty TF, McFadyen I, Scarborough M, et al. Fracture-related infection: A consensus on definition from an international expert group. Injury. 2018;49(3):505–10.

273. Osmon DR, Berbari EF, Berendt AR, Lew D, Zimmerli W, Steckelberg JM, et al. Diagnosis and Management of Prosthetic Joint Infection : Clinical Practice Guidelines by the Infectious Diseases Society of America a. Clin Infect Dis. 2013;56:1–25.

274. Tetreault MW, Wetters NG, Aggarwal VK, Moric M, Segreti J, Huddleston JI, et al.

Should draining wounds and sinuses associated with hip and knee arthroplasties be cultured? J Arthroplasty. 2013;28(8 SUPPL):133–6.

275. Hoekstra M, Veltman ES, Nurmohamed RFRHA, Dijk B Van, Rentenaar RJ, Vogely HC, et al. Sonication Leads to Clinically Relevant Changes in Treatment of Periprosthetic Hip or Knee Joint Infection. J Bone Jt Infect. 2020;5(3):128–32.

276. Hughes HC, Newnham R, Athanasou N, Atkins BL, Bejon P, Bowler ICJW. Microbiological diagnosis of prosthetic joint infections: A prospective evaluation of four bacterial culture media in the routine laboratory. Clin Microbiol Infect. 2011;17(10):1528–30.

277. Schwotzer N, Wahl P, Fracheboud D, Gautier E, Chuard C. Optimal culture incubation time in orthopedic device-associated infections: A retrospective analysis of prolonged 14-day incubation. J Clin Microbiol. 2014;52(1):61–6.

278. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6:1–16.

279. British Society for Antimicrobial Chemotherapy. BSAC Methods for Antimicrobial Susceptibility Testing. Vol. Version 44. 2015.

280. British Society for Antimicrobial Chemotherapy. Standing Committee on Susceptibility Testing Version 14.0. 2015;Version 14.

281. Winslow C-EA, Winslow AR. The Systematic Relationships of the Coccaceae: With a Discussion of the Principles of Bacterial Classification. Wiley; 1908.

282. Zhong C, Nelson M, Cao G, Sadowsky MJ, Yan T. Complete Genome Sequence of Staphylococcus epidermidis ATCC 12228 Chromosome and Plasmids, Generated by Long-Read Sequencing. Genome Announc. 2017;5(36):4–5.

283. Lunter G, Goodson M. Stampy : A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. 2011;21(3):936–9.

284. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18(5):821–9.

285. Golubchik T, Batty EM, Miller RR, Farr H, Young BC, Larner-Svensson H, et al. Within-Host Evolution of Staphylococcus aureus during Asymptomatic Carriage. PLoS One. 2013;8(5):1–14.

286. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

287. Agostino M. Introduction to the BLAST Suite and BLASTN. Pract Bioinforma. 2012;26.

288. Staphylococcus epidermidis MLST Home Page [Internet]. [cited 2018 Nov 26]. Available from: https://pubmlst.org/sepidermidis/

289. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST + : architecture and applications. BMC Bioinformatics. 2009;9:1–9.

290. Sanner MF, Jolla L. Python: a programming language for software integration and development. J Mol Graph Model. 1999;17(1):57–61.

291. Mason A, Foster D, Bradley P, Golubchik T, Doumith M, Gordon NC, et al. Accuracy of different bioinformatics methods in detecting antibiotic resistance and virulence factors from staphylococcus aureus whole-genome sequences. J Clin Microbiol. 2018;56(9):1815–7.

292. Fredheim EGA, Flaegstad T, Askarian F, Klingenberg C, Flaegstad T, Askarian F, et al. Colonisation and interaction between S. epidermidis and S. aureus in the nose and throat of healthy adolescents. Eur J Clin Microbiol Infect Dis. 2014;34(1):123–9.

293. Roth R, James WD. Microbial Ecology Of The Skin. Annu Rev Microbiol. 1988;42(1):441–64.

294. Kloos WE, Musselwhite MS. Distribution and persistence of Staphylococcus and Micrococcus species and other aerobic bacteria on human skin. Appl Microbiol. 1975;30(3):381–5.

295. Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. Nat Rev Microbiol. 2018;16(3):143.

296. Otto M. Staphylococcus colonization of the skin and antimicrobial peptides. Expert Rev Dermatol. 2010;5(2):183–95.

297. pubmlst.org. S. epidermidis MLST profiles [Internet]. [cited 2019 Nov 11]. Available from: https://pubmlst.org/bigsdb?db=pubmlst_sepidermidis_seqdef&page=downloadProfiles&scheme_id=1

298. Larson EL, McGinley KJ, Foglia AR, Talbot GH, Leyden JJ. Composition and antimicrobic resistance of skin flora in hospitalized and healthy adults. J Clin Microbiol. 1986;23(3):604–8.

299. Terpstra S, Noordhoek GT, Voesten HGJ, Hendriks B, Degener JE. Rapid emergence of resistant coagulase-negative staphylococci on the skin after antibiotic prophylaxis. J Hosp Infect. 1999;43(3):195–202.

300. Zhang L, Thomas JC, Miragaia M, Bouchami O, Chaves F, d'Azevedo PA, et al. Multilocus Sequence Typing and Further Genetic Characterization of the Enigmatic Pathogen, Staphylococcus hominis. PLoS One. 2013;8(6).

301. Widerstrom M, Wistrom J, Sjostedt A, Monsen T, Widerström M, Wiström J, et al. Coagulase-negative staphylococci: Update on the molecular epidemiology and clinical presentation, with a focus on Staphylococcus epidermidis and Staphylococcus saprophyticus. Eur J Clin Microbiol Infect Dis. 2012 Jan;31(1):7–20.

302. Hedin G. A comparison of methods to determine whether clinical isolates of Staphylococcus epidermidis from the same patient are related. J Hosp Infect. 1996;34(1):31–42.

303. Begović J, Jovčić B, Papić-Obradović M, Veljović K, Lukić J, Kojić M, et al. Genotypic diversity and virulent factors of Staphylococcus epidermidis isolated from human breast milk. Microbiol Res. 2013;168(2):77–83.

304. Paterson GK, Harrison EM, Murray GGR, Welch JJ, Warland JH, Holden MTG, et al. Capturing the cloud of diversity reveals complexity and heterogeneity of MRSA carriage, infection and transmission. Nat Commun. 2015;6:1–10.

305. Arslan S, Ozkardes F. Slime production and antibiotic susceptibility in staphylococci isolated from clinical samples. Mem Inst Oswaldo Cruz. 2007;102(1):29–33.

306. Hübner J, Kropec A. Cross infections due to coagulase-negative staphylococci in high-risk patients. Zentralbl Bakteriol. 1995;283(2):169–74.

307. Foster T. Staphylococcus. In: Medical Microbiology. 4th edition. 1996. Chapter 12. Galveston (TX): University of Texas Medical Branch at Galveston;

308. Young BC, Golubchik T, Batty EM, Fung R, Larner-Svensson H, Votintseva AA, et al. Evolutionary dynamics of Staphylococcus aureus during progression from carriage to disease. ProcNatlAcadSciUSA. 2012;109(12):4550–5.

309. Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carriço JA. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. BMC Bioinformatics. 2012;13(1):87.

310. Francisco AP, Bugalho M, Ramirez M, Carriço JA. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. 2009;15:1–15.

311. Didelot X, Wilson DJ. ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. PLoS Comput Biol. 2015;11(2):1–18.

312. Longueville J-E, Lefort V, Gascuel O. SMS: Smart Model Selection in PhyML. Mol Biol Evol. 2017 May 11;34(9):2422–4.

313. Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. Genetics. 1992;132(2):583–9.

314. Krause DJ, Whitaker RJ. Inferring speciation processes from patterns of natural variation in microbial genomes. Syst Biol. 2015;64(6):926–35.

315. Export dataset - Staphylococcus epidermidis isolates [Internet]. [cited 2018 Nov 21]. Available from: https://pubmlst.org/bigsdb?page=plugin&name=Export&db=pubmlst_sepidermidis_isolates

316. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5.

317. Price JR, Cole K, Bexley A, Kostiou V, Eyre DW, Golubchik T, et al. Transmission of Staphylococcus aureus between health-care workers, the environment, and patients in an intensive care unit: a longitudinal cohort study based on whole-genome sequencing. Lancet Infect Dis. 2016;17(2):207–14.

318. Young BC, Wu C-HH, Gordon NC, Cole K, Price JR, Liu E, et al. Severe infections emerge from commensal bacteria by adaptive evolution. Elife. 2017;6:1–25.

319. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289–90.

320. R Core Team, R: A language and environment for statistical computing. 2013;

321. Zuckerkandl E, Pauling L. Molecules as documents of evolutionary history. J Theor Biol. 1965;8(2):357–66.

322. Ho S. The molecular clock and estimating species divergence. Nat Educ. 2008;1(1):1–2.

323. Reis M, Donoghue PCJ, Yang Z. Bayesian molecular clock dating of species divergences in the genomics era. Nat Publ Gr. 2015;

324. Zhou J, Teo Y. Estimating time to the most recent common ancestor (TMRCA): comparison and application of eight methods. Eur J Hum Genet. 2015/12/16. 2016 Aug;24(8):1195–201.

325. Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson J. Bayesian inference of ancestral dates on bacterial phylogenetic trees. Nucleic Acids Res. 2018;46(22).

326. Cavanagh JP, Hjerde E, Holden MTG, Kahlke T, Klingenberg C, Flaegstad T, et al. Whole-genome sequencing reveals clonal expansion of multiresistant Staphylococcus

haemolyticus in European hospitals. J Antimicrob Chemother. 2014;69(11):2920–7.

327. Hal SJ Van, Ip CLC, Ansari MA, Wilson DJ, Espedido BA, Jensen SO, et al. Evolutionary dynamics of Enterococcus faecium reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. Microb Genomics. 2016;

328. Price JR, Golubchik T, Cole K, Wilson DJ, Crook DW, Thwaites GE, et al. Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of staphylococcus aureus in an intensive care unit. Clin Infect Dis. 2014;58(5):609–18.

329. Cherifi S, Byl B, Deplano A, Nagant C, Nonhoff C, Denis O, et al. Genetic characteristics and antimicrobial resistance of Staphylococcus epidermidis isolates from patients with catheter-related bloodstream infections and from colonized healthcare workers in a Belgian hospital. Ann Clin Microbiol Antimicrob. 2014;13(1):20.

330. Eyre DW, Golubchik T, Gordon NC, Bowden R, Piazza P, Batty EM, et al. A pilot study of rapid benchtop sequencing of Staphylococcus aureus and Clostridium difficile for outbreak detection and surveillance. BMJ Open. 2012;2(3):e001124–e001124.

331. Haunreiter VD, Boumasmoud M, Häffner N, Wip D, Leimer N, Rachmühl C, et al. In-host evolution of Staphylococcus epidermidis in a pacemaker-associated endocarditis resulting in increased antibiotic tolerance. Nat Commun. 2019;10(1):1149.

332. Frei R, Dangel M, Jakob M, Balmelli C, Schaefer DJ, Weisser M, et al. Validation of a treatment algorithm for orthopaedic implant-related infections with device-retention — results from a prospective observational cohort study. Clin Microbiol Infect. 2016;22(5):457.e1-457.e9.

333. Widmer AF. New Developments in Diagnosis and Treatment of Infection in Orthopedic Implants. Clin Infect Dis. 2001;33(s2):S94–106.

334. Bozic KJ, Kurtz SM, Lau E, Ong K, Vail TP, Berry DJ. The Epidemiology of Revision Total Hip Arthroplasty in the United States. J Bone Jt Surg. 2009;91(1):128–33.

335. Sharkey PF, Hozack WJ, Rothman RH, Shastri S, Jacoby SM. Why are total knee arthroplasties failing today? Clin Orthop Relat Res. 2002;(404):7–13.

336. Ulrich SD, Seyler TM, Bennett D, Delanois RE, Saleh KJ, Thongtrangan I, et al. Total hip arthroplasties: What are the reasons for revision? Int Orthop. 2008;32(5):597–604.

337. Klouche S, Sariali E, Mamoudy P. Total hip arthroplasty revision due to infection: a cost analysis approach. Orthop Traumatol Surg Res. 2010;96(2):124–32.

338. Shapiro JA, Nguyen VL, Chamberlain NR. Evidence for persisters in Staphylococcus epidermidis RP62a planktonic cultures and biofilms. J Med Microbiol. 2011;60(7):950–60.

339. James PJ, Butcher IA, Gardner ER, Hamblen DL. Methicillin-resistant Staphylococcus epidermidis in infection of hip arthroplasties. Bone Joint J. 1994;76(5):725–7.

340. Tarabichi M, Shohat N, Goswami K, Alvand A, Silibovsky R, Belden K, et al. Diagnosis of Periprosthetic Joint Infection : The Potential of Next-Generation Sequencing. J Bone Jt Surg. 2018;100:147–54.

341. Ruppé E, Lazarevic V, Girard M, Mouton W, Ferry T, Schrenzel J, et al. Clinical metagenomics of bone and joint infections: a proof of concept study. Nat Sci Reports. 2017;7(1):1–12.

342. Street TL, Sanderson ND, Atkins BL, Brent AJ, Foster D, Mcnally MA, et al. Molecular

diagnosis of orthopaedic device infection direct from sonication fluid by metagenomic sequencing. J Clin Microbiol. 2017;55(8):2334–47.

343. Huang Z, Li W, Fang X, Lin J, Zhang W, Lee GC, et al. Metagenomic next-generation sequencing of synovial fluid demonstrates high accuracy in prosthetic joint infection diagnostics. Bone Jt Res. 2020;9(7):440–9.

344. Ivy MI, Thoendel MJ, Jeraldo PR, Greenwood-Quaintance KE, Hanssen AD, Abdel MP, et al. Direct detection and identification of prosthetic joint infection pathogens in synovial fluid by metagenomic shotgun sequencing. J Clin Microbiol. 2018;56(9).

345. Wang C, Huang Z, Li W, Fang X, Zhang W. Can metagenomic next-generation sequencing identify the pathogens responsible for culture-negative prosthetic joint infection? BMC Infect Dis. 2020;20(1):1–7.

346. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. Within-host evolution of bacterial pathogens. Nat Rev Microbiol. 2016;14(3):150–62.

347. Vanhegan IS, Malik AK, Jayakumar P, Ul Islam S, Haddad FS. A financial analysis of revision hip Arthroplasty. J Bone Jt Surg. 2012;94 B(5):619–23.

348. The Health Service (Control of Patient Information) Regulations 2002 No. 1438. National Health Service, England and Wales. p. 1-8.

349. Mann Whitney U Test (Wilcoxon Rank Sum Test) [Internet]. Boston University. [cited 2018 Apr 25]. Available from: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/bs704_nonparametric4.html

350. Nachar N. The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. Tutor Quant Methods Psychol. 2008;4(1):13–20.

351. Parvizi J, Zmistowski B, Berbari EF, Bauer TW, Springer BD, Della Valle CJ, et al. New definition for periprosthetic joint infection: From the workgroup of the musculoskeletal infection society. Clin Orthop Relat Res. 2011;469(11):2992–4.

352. Makhlin J, Kofman T, Borovok I, Kohler C, Engelmann S, Cohen G, et al. Staphylococcus aureus ArcR controls expression of the arginine deiminase operon. J Bacteriol. 2007;189(16):5976–86.

353. Wisplinghoff H, Rosato AE, Enright MC, Noto M, Craig W, Archer GL. Related Clones Containing SCC mec Type IV Predominate among Clinically Significant Staphylococcus epidermidis Isolates Related Clones Containing SCCmec Type IV Predominate among Clinically Significant Staphylococcus epidermidis Isolates. Antimicrob Agents Chemother. 2003;47(11):3574–9.

354. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

355. Carattoli A, Zankari E, Garciá-Fernández A, Larsen MV, Lund O, Villa L, et al. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895–903.

356. Lenguerrand E, Whitehouse MR, Beswick AD, Kunutsor SK, Foguet P, Porter M, et al. Risk factors associated with revision for prosthetic joint infection following knee replacement: an observational cohort study from England and Wales. Lancet Infect Dis. 2019;19(6):589–600.

357. Büttner H, Mack D, Rohde H. Structural basis of Staphylococcus epidermidis biofilm formation: mechanisms and molecular interactions. Front Cell InfectMicrobiol.

2015;5(14):1–15.

358. Xu Z, Misra R, Jamrozy D, Paterson GK, Cutler RR, Holmes MA, et al. Whole Genome Sequence and Comparative Genomics Analysis of Multi-drug Resistant Environmental Staphylococcus epidermidis ST59. G3 Genes|Genomes|Genetics. 2018;8(July):2225–30.

359. Chaudhry V, Patil PB. Genomic investigation reveals evolution and lifestyle adaptation of endophytic Staphylococcus epidermidis. Sci Rep. 2016;6(October 2015):19263.

360. Silva NCC, Guimarães FF, Manzi MDP, Gómez-Sanz E, Gómez P, Araújo JP, et al. Characterization of methicillin-resistant coagulase-negative staphylococci in milk from cows with mastitis in Brazil. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2014;106(2):227–33.

361. Fajardo-Cavazos P, Nicholson WL. Cultivation of Staphylococcus epidermidis in the human spaceflight environment leads to alterations in the frequency and spectrum of spontaneous rifampicin-resistance mutations in the rpoB gene. Front Microbiol. 2016;7(JUN):1–10.

362. Lee JYHH, Monk IR, Gonçalves da Silva A, Seemann T, Chua KYLL, Kearns A, et al. Global spread of three multidrug-resistant lineages of Staphylococcus epidermidis. Nat Microbiol. 2018;3(10):1175–85.

363. Ashimoto A, Hamada T, Adachi A, Tanigawa T, Tanaka Y. Molecular epidemiology of Staphylococcus spp. contamination in the ward environment: study on mecA and femA genes in methicillin-resistant strains. J Japanese Assoc Infect Dis. 1995;69(1):15–20.

364. de Almeida LM, Lincopan N, de Araújo MRE, Mamizuka EM. Dissemination of the linezolid-resistant Staphylococcus epidermidis clone ST2 exhibiting the G2576t mutation in the 23s rRNA gene in a tertiary-care hospital, Brazil. J Antimicrob Chemother. 2012;67(3):768–9.

365. Ramos-Trujillo E, Pérez-Roth E, Méndez-Alvarez S, Claverie-Martín F. Multiplex PCR for simultaneous detection of enterococcal genes vanA and vanB and staphylococcal genes mecA, ileS-2 and femB. Int Microbiol. 2003;6(2):113–5.

366. Martineau F, Picard FJ, Grenier L, Roy PH, Ouellette M, Bergeron MG. Multiplex PCR assays for the detection of clinically relevant antibiotic resistance genes in staphylococci isolated from patients infected after cardiac surgery. The ESPRIT Trial. JAntimicrobChemother. 2000;46(0305-7453 SB-IM):527–34.

367. Zhu L-X, Zhang Z-W, Wang C, Yang H-W, Jiang D, Zhang Q, et al. Use of a DNA microarray for simultaneous detection of antibiotic resistance genes among staphylococcal clinical isolates. J Clin Microbiol. 2007;45(11):3514–21.

368. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Philippon A, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020 Dec 1;75(12):3491–500.

369. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J Antimicrob Chemother. 2017 Oct 1;72(10):2764–8.

370. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45(D1):D566–73.

371. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb genomics. 2017;3(10):1–11.

372. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother. 2014;58(1):212–20.

373. Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, et al. Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. Nat Commun. 2015;6(1):1–15.

374. Pankhurst LJ, del Ojo Elias C, Votintseva AA, Walker TM, Cole K, Davies J, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: A prospective study. Lancet Respir Med. 2016;4(1):49–58.

375. Eyre DW, De Silva D, Cole K, Peters J, Cole MJ, Grad YH, et al. WGS to predict antibiotic MICs for Neisseria gonorrhoeae. J Antimicrob Chemother. 2017;72(2):1937–47.

376. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, et al. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J Antimicrob Chemother. 2013;68(10):2234–44.

377. Deng X, Memari N, Teatero S, Athey T, Isabel M, Mazzulli T, et al. Whole-genome Sequencing for Surveillance of Invasive Pneumococcal Diseases in Ontario, Canada: Rapid Prediction of Genotype, Antibiotic Resistance and Characterization of Emerging Serotype 22F. Front Microbiol. 2016 Dec 27;7(December):2099.

378. Lauener FN, Imkamp F, Lehours P, Buissonnière A, Benejat L, Zbinden R, et al. Genetic determinants and prediction of antibiotic resistance phenotypes in Helicobacter pylori. J Clin Med. 2019;8(1):53.

379. Sadouki Z, Day MR, Doumith M, Chattaway MA, Dallman TJ, Hopkins KL, et al. Comparison of phenotypic and WGS-derived antimicrobial resistance profiles of Shigella sonnei isolated from cases of diarrhoeal disease in England and Wales, 2015. J Antimicrob Chemother. 2017 Sep 1;72(9):2496–502.

380. Nguyen M, Brettin T, Long SW, Musser JM, Randall J, Olson R, et al. Developing an in silico minimum inhibitory concentration panel test for Klebsiella pneumoniae. Sci Rep. 2018;8(1):421.

381. Ruesen C, Riza AL, Florescu A, Chaidir L, Editoiu C, Aalders N, et al. Linking minimum inhibitory concentrations to whole genome sequence-predicted drug resistance in Mycobacterium tuberculosis strains from Romania. Sci Rep. 2018;(February):1–8.

382. Li Y, Metcalf BJ, Chochua S, Li Z, Jr REG, Walker H, et al. Validation of β-lactam minimum inhibitory concentration predictions for pneumococcal isolates with newly encountered penicillin binding protein (PBP) sequences. BMC Genomics. 2017;18(1):621.

383. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007 Nov;23(21):2947–8.

384. Antonio M, McFerran N, Pallen MJ. Mutations affecting the Rossman fold of isoleucyl-tRNA synthetase are correlated with low-level mupirocin resistance in Staphylococcus aureus. Antimicrob Agents Chemother. 2002;46(2):438–42.

385. Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. Trends Genet. 2014;30(9):401–7.

386. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agers?? Y, Lund O, et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother. 2013;68(4):771–7.

387. Didelot X, Bowden R, Wilson DJ, Peto TE a., Crook DW. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet. 2012;13(9):601–12.

388. Armstrong GL, MacCannell DR, Taylor J, Carleton HA, Neuhaus EB, Bradbury RS, et al. Pathogen genomics in public health. N Engl J Med. 2019;381(26):2569–80.

389. McDermott PF, Tyson GH, Kabera C, Chen Y, Li C, Folster JP, et al. Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal Salmonella. Antimicrob Agents Chemother. 2016;60(9):5515–20.

390. Zhao S, Tyson GH, Chen Y, Li C, Mukherjee S, Young S, et al. Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. Appl Environ Microbiol. 2016;82(2):459–66.

391. Clausen PTLC, Zankari E, Aarestrup FM, Lund O. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. J Antimicrob Chemother. 2016;71(9):2484–8.

392. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry and FDA. Class II Special Controls Guidance Document : Antimicrobial Susceptibility Test Systems. 2009;1–42.

393. Aubry B, Lemarié C, Chenouard R, Kempf M, Eveillard M, Pailhoriès H. Performance of penicillinase detection tests in Staphylococcus epidermidis: comparison of different phenotypic methods. BMC Microbiol. 2020;20(1):1–8.

394. Committee TE, Testing AS, Changes N, Pseudomonas E. European Committee on Antimicrobial Susceptibility Testing Breakpoint tables for interpretation of MICs and zone diameters European Committee on Antimicrobial Susceptibility Testing Breakpoint tables for interpretation of MICs and zone diameters. http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_50_Breakpoint_Table_01.pdf. 2021;0–77.

395. Peacock SJ, Paterson GK. Mechanisms of Methicillin Resistance in *Staphylococcus aureus*. Annu Rev Biochem. 2015;84(1):577–601.

396. Flamm RK, Mendes RE, Hogan PA, Streit JM, Ross JE, Jonesa RN. Linezolid surveillance results for the United States (LEADER surveillance program 2014). Antimicrob Agents Chemother. 2016;60(4):2273–80.

397. Decousser JW, Desroches M, Bourgeois-Nicolaos N, Potier J, Jehl F, Lina G, et al. Susceptibility trends including emergence of linezolid resistance among coagulase-negative staphylococci and meticillin-resistant Staphylococcus aureus from invasive infections. Int J Antimicrob Agents. 2015;46(6):622–30.

398. Tevell S, Claesson C, Hellmark B, Söderquist B, Nilsdotter-Augustinsson Å. Heterogeneous glycopeptide intermediate Staphylococcus epidermidis isolated from prosthetic joint infections. Eur J Clin Microbiol Infect Dis. 2014;33(6):911–7.

399. Depardieu F, Podglajen I, Leclercq R, Collatz E, Courvalin P. Modes and modulations of antibiotic resistance gene expression. Clin Microbiol Rev. 2007;20(1):79–114.

400. Suzuki S, Horinouchi T, Furusawa C. Prediction of antibiotic resistance by gene expression profiles. Nat Commun. 2014;5.

401. Sanderson ND, Swann J, Barker L, Kavanagh J, Hoosdally S, Crook D, et al. High

precision Neisseria gonorrhoeae variant and antimicrobial resistance calling from metagenomic Nanopore sequencing. Genome Res. 2020;30(9):1354–63.

402. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640–4.

403. Widerström M, McCullough CA, Coombs GW, Monsen T, Christiansen KJ. A multidrug-resistant Staphylococcus epidermidis clone (ST2) is an ongoing cause of hospital-acquired infection in a Western Australian Hospital. J Clin Microbiol. 2012;50(6):2147–51.

404. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect. 2012;18(3):268–81.

405. Perry KI, Hanssen AD. Orthopaedic infection: prevention and diagnosis. JAAOS-Journal Am Acad Orthop Surg. 2017;25(suppl 1):S4–6.

406. Miller LS, Fowler VG, Shukla SK, Rose WE, Proctor RA. Development of a vaccine against Staphylococcus aureus invasive infections: Evidence based on human immunity, genetics and bacterial evasion mechanisms. Vol. 44, FEMS Microbiology Reviews. 2019. 123–153 p.

407. Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics. Public Health England. 2018.

408. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 6.0, 2016. European Committee on Antimicrobial Susceptibility Testing Breakpoint tables for interpretation of MICs and zone diameters Version 6.0, valid from 2016-01-01. 2015. p. 0–77.