Susana Maria Pereira Brás

The adaptation of *Staphylococcus epidermidis* commensal and clinical isolates to human blood: the search for molecular diagnostics markers
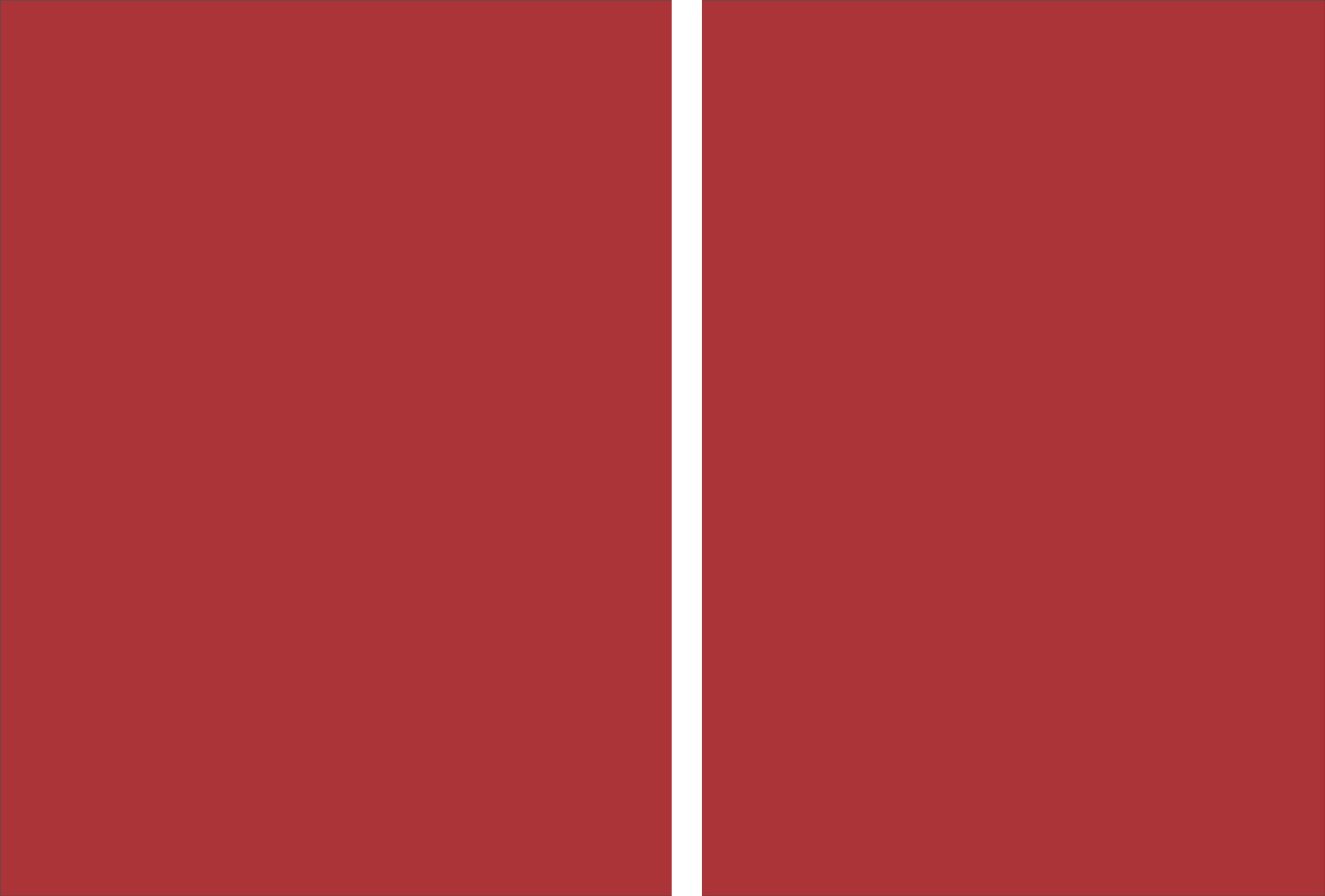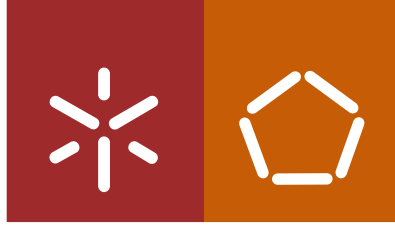
janeiro de 2020

Universidade do Minho
Escola de Engenharia

Susana Maria Pereira Brás

The adaptation of *Staphylococcus epidermidis* commensal and clinical isolates to human blood: the search for molecular diagnostics markers

Tese de Doutoramento
Doutoramento em Engenharia Química e Biológica

Trabalho efetuado sob a orientação do
Doutor Nuno Cerca
e da
Doutora Ângela França

janeiro de 2020

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho

Acknowledgements

A realização desta desafiante etapa de vida, recheada de momentos bons e momentos maus, só foi possível devido à presença de pessoas a quem dedicarei, de todo o coração, algumas palavras sinceras de agradecimento.

Ao meu orientador **Doutor Nuno Cerca**, agradeço a oportunidade, o apoio, a amizade, a partilha de conhecimentos e de sugestões. À minha co-orientadora **Doutora Ângela França** todas as suas sugestões e pela partilha de conhecimentos.

A todos os meus colegas e amigos do grupo "**Cerca *et al***" agradeço o apoio, os sorrisos, a partilha, os conselhos e a amizade, em especial aos meus companheiros de percurso, **Fernando** e **Joana**.

A todas as pessoas que de alguma forma tiveram um contributo importante ao longo desta longa etapa deixo aqui o meu profundo agradecimento.

A toda a minha família e amigos agradeço todo o apoio e amor dados.

E por fim, às pessoas mais importantes da minha vida e a quem dedico. Ao meu **pai** por ser o pilar da minha vida (mesmo no alto dos meus 40 anos), pela inspiração e por todo o apoio. Ao meu companheiro de vida, o **Daniel**, pelo amor, apoio e a ajuda para concretizar esta etapa. E à minha linda e adorável filha **Júlia**, por todo o seu amor incondicional.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

A adaptação de isolados clínicos e comensais de *Staphylococcus epidermidis* ao sangue humano: a procura de marcadores moleculares de diagnóstico.

Resumo

Staphylococcus epidermidis é um habitante normal de pele e mucosas humanas saudáveis que está associado a infeções nosocomiais, principalmente em pacientes imunocomprometidos, contribuindo para um aumento considerável nos custos na saúde, na morbilidade e na mortalidade. Isto está relacionado principalmente com a crescente utilização de dispositivos médicos e com a capacidade de *S. epidermidis* de formar biofilmes nesses dispositivos. Uma questão importante que precisa ser esclarecida prende-se com a dificuldade de diagnosticar as infeções causadas por *S. epidermidis*, uma vez que ocorre com facilidade contaminação das hemoculturas devido à natureza onipresente na pele humana e a dificuldade em diferenciar isolados comensais da pele de isolados invasivos que causam doença. Assim, o objetivo desta tese foi tentar encontrar marcadores moleculares baseados na expressão de RNA de forma a diferenciar os isolados comensais dos isolados clínicos, e assim, contribuir para o desenvolvimento de uma nova ferramenta de diagnóstico que permita discriminar os isolados que contaminam as hemoculturas dos isolados associados a infeções. Para atingir esse objetivo, começou por analisar-se o transcriptoma de três isolados clínicos e de três isolados comensais, após duas horas de incubação em sangue humano, tendo sido destacado cinco potenciais genes que permitiram a diferenciação destes seis isolados. Infelizmente, quando se tentou validar esses resultados numa coleção de 56 isolados, o potencial discriminatório desses genes foi perdido. Por outro lado, estes resultados destacaram a grande versatilidade metabólica dos isolados que ocupam os diferentes nichos, o que reforça a ideia que *S. epidermidis*, é de facto, um patógeno oportunista. Essa segunda hipótese foi confirmada através da avaliação da capacidade de sobrevivência em sangue humano, onde foi demonstrado que todos os isolados testados tinham a mesma capacidade de adaptação ao sangue humano, o que sugere que a capacidade de *S. epidermidis* causar infeção depende da oportunidade de penetrar as camadas externas protetoras do hospedeiro e da sua capacidade de sobreviver à ação antimicrobiana do sangue humano.

Palavras-chave: Diagnóstico, Discriminação, Sangue humano, *Staphylococcus epidermidis* transcriptoma.

The adaptation of *Staphylococcus epidermidis* commensal and clinical isolates to human blood: the search for molecular diagnostics markers

Abstract

Staphylococcus epidermidis is a normal inhabitant of healthy human skin and mucosae that originate important nosocomial infections, particularly in immunocompromised patients, contributing to a considerable increase in healthcare costs, morbidity, and mortality. This is mainly related to the increasing utilization of medical devices and the capacity of *S. epidermidis* to form biofilms on such devices. An important issue that needs to be addressed is related to the difficulty to accurately diagnose *S. epidermidis* infections, which consequently results in prolonged hospital stays, inappropriate treatment, and further testing. This occurs due to the easiness of contamination of blood samples collected for diagnosis, due to the ubiquitous nature of this species in the human skin, and the difficulty in differentiating skin commensal isolates from invasive isolates that cause disease. The purpose of this thesis was to identify RNA-based molecular markers that could be able to differentiate clinical from commensal isolates and, ultimately, to contribute to the development of a novel diagnostic tool that could accurately discriminate strains contaminating blood cultures from infection-associated strains. The analysis of the transcriptome of three clinical and three commensal isolates, after two hours of incubation in human blood, highlighted five potential markers. Unfortunately, when we tried to validate these findings with a collection of 56 isolates, the discriminatory potential of these genes was lost. Conversely, these results highlighted the great metabolic versatility of *S. epidermidis* isolates from different niches, which reinforces the idea that *S. epidermidis* is, in fact, an opportunistic pathogen. We further confirmed this hypothesis by assessing bacterial survivability in human blood, where we demonstrated that both clinical and commensal isolates had similar ability to adapt to the host microenvironment, in this case, the human blood. Overall, in our experimental model, the isolates from hospital and community settings displayed the same ability to survive and proliferate in human blood further suggesting that *S. epidermidis* ability to cause infection mostly relies on the opportunity to breach the protective layers of the host and its ability to evade and proliferate in the immunocompromised host.

Keywords: Diagnosis, Discrimination, Human blood, *Staphylococcus epidermidis*, Transcriptome

Table of Contents

List of abbreviations

Aap	Accumulation-associated protein
Agr	Accessory gene regulator
ACME	Arginine catabolic mobile element
AMP	Antimicrobial peptides
Bhp	Homologue to the biofilm-associated protein of *Staphylococcus aureus*
bp	Base pair
cDNA	Complementary deoxyribonucleic acid
CDC	Centers for Disease Control and Prevention
CFU	Colony-forming unit
CoNS	Coagulase-negative staphylococci
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetraacetic acid
Embp	Extracellular matrix-binding protein
GC	Genetic clusters
ica	Intercellular adhesion operon
LTA	Lipoteichoic acid
mRNA	Messenger ribonucleic acid
MRSA	Methicillin resistant *Staphylococcus aureus*
MLST	Multilocus sequence typing
NaCl	Sodium chloride
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI	Propidium iodide
PIA/PNAG	Polysaccharide intercellular adhesin/ poly-N-acetylglucosamine
PMNs	Polymorphonuclear leukocytes
PSM	Phenol-soluble modulin
qPCR	Quantitative PCR
RNA	Ribonucleic acid
RNA-seq	RNA sequencing

RPKM	Reads per kilo base per million mapped reads
ROS	Reactive oxygen species
RT	Room temperature
Sbp	Small basic protein
TAE	Tris–acetate–EDTA
TLR	Toll-like receptor
TSA	Tryptic soy agar
TSB	Tryptic soy broth
TSBG	TSB supplemented with 0.4% (w/v) glucose
SCC*mec*	Staphylococcal cassette chromosome *mec*
WGS	Whole genome sequence

List of figures

List of tables

Scientific outputs derived from this thesis

Oral communication

Brás S, Cerca N, França A. "On the old story of the accidental pathogen: can transcriptomics clarify how *S. epidermidis* becomes virulent?" Congress of Microbiology and Biotechnology 2019, Dec 5-7, Coimbra, Portugal.

Poster communication at international scientific conferences

Brás S, França A, Cerca N. Developing an affordable but reliable human blood *ex vivo* model to analyze gene expression by *Staphylococcus epidermidis.* Congress of Microbiology and Biotechnology 2017, Dec 7-9, Porto, Portugal.

Brás S, França A, Cerca N. "Probing for RNA sequencing data in search for possible diagnostic markers for S*taphylococcus epidermidis* infections". International Symposium of Staphylococci and Staphylococcal infections 2018, August 23-26, Copenhagen, Denmark.

Ao meu pai

Aos meus amores

CHAPTER 1

Introduction

SUMMARY

This chapter provides a brief outline of the thesis. The background, research questions, hypothesis, aims, and significance are presented here.

1.1 BACKGROUND

Staphylococcus epidermidis is a normal inhabitant of healthy human skin and mucosae that can originate important infections if it gains access into the bloodstream [1]. The increasing use of medical devices for diagnosis and therapeutic procedures provides the opportunity for *S. epidermidis* to colonize and invade its host [2]. Nowadays, *S. epidermidis* is recognized as the most frequent cause of catheter-related bloodstream infections [3]. These infections can be very difficult to diagnose which also contributes to difficulties in treatment. Therefore, bloodstream infections have a huge impact on healthcare costs, morbidity, and mortality. An accurate and fast diagnosis is the key to reduce the clinical and economic impact of *S. epidermidis* infections.

Infections caused by *S. epidermidis* are difficult to diagnose, due to the difficulty in differentiating bacteremia from culture contamination [4]. The absence of determinants of pathogenicity and the fact *S. epidermidis* bacteremia presents as an indolent infection, the discrimination of true bacteremia from blood culture contamination is problematic. In fact, blood culture contamination represents an ongoing problem that often leads to misdiagnosis, which consequently results in a significant increase in healthcare costs due to prolonged hospital stays, treatment, and further testing [5]. In the last years, several studies have attempted to identify potential markers that could allow the discrimination of isolates that cause infection from culture contamination (presumably caused by commensal isolates). These attempts have focused on phenotypic and genomic approaches, however, no marker was found that was able to differentiate isolates that cause infection from contaminating isolates.

In this study, the strategy to discriminate between clinical and commensal isolates will be assessed based on the different transcriptomic profiles. The transcriptome of three commensal and three clinical isolates, incubated in human blood, will be compared to highlight molecular markers that can to discriminate between the two groups. The transcription levels of these putative molecular markers will be then validated using a worldwide collection of 56 *S. epidermidis* strains. A comparison of gene expression profiles between clinical and commensal isolates is expected to reveal genes that may prove useful as diagnostic markers of *S. epidermidis* bacteremia.

1.2 RESEARCH QUESTIONS

The following questions will be addressed in this thesis:

1. Can transcriptome sequencing data provide the identification of molecular diagnostic markers to differentiate clinical isolates from commensal isolates?
2. Do clinical and commensal isolates have equal capacity to survive in human blood?
3. Can we reduce the amount of human blood used in an *ex vivo* model without compromising the analysis of *S. epidermidis* gene expression?

Answers to these research questions are expected to provide new knowledge that could assist the diagnosis of *S. epidermidis* infections.

1.3 HYPOTHESIS AND AIMS

1.3.1 Hypothesis

Differential expression of specific genes determines the ability of *S. epidermidis* to adapt to an environment like human blood and, eventually cause an infection. These differences can be used as a diagnostic marker to detect invasive *S. epidermidis* strains.

1.3.2 Aims

The overall aim of this thesis was to assess if transcriptomic data could be used as a diagnostic tool to differentiate *S. epidermidis* true bacteremia from culture contamination. A human blood *ex vivo* model was used to determine the adaptations of clinical and commensal isolates in the human blood. To achieve this ultimate goal, several specific aims were devised:

Aim 1: To perform an *in silico* comparative transcriptomic analysis between clinical and commensal isolates using RNA sequencing data, to identify and select possible molecular markers that can help to diagnose *S. epidermidis* bacteremia.

Aim 2: To investigate whether the transcriptional levels of the selected markers can discriminate clinical isolates from commensal isolates using a collection of *S. epidermidis* isolates.

Aim 3: To assess key parameters (volume of blood, the concentration of bacteria, time of incubation) that are expected to influence *S. epidermidis* gene expression when using human blood as an *ex vivo* model.

1.4 SIGNIFICANCE

The infections caused by *S. epidermidis* are indolent and the detection of *S. epidermidis* may signify bacteremia or culture contamination [4]. It is noteworthy that blood culture contamination represents an ongoing problem that often leads to misdiagnosis and precipitate unnecessary antibiotic exposure and hospitalization [6]. Over the years, a significant proportion of research has attempted to find diagnostic markers with no success so far. Apparently, no genetic differences truly discriminate clinical from commensal isolates.

To date, no RNA sequencing studies have investigated transcriptomic profiles to characterize invasive *S. epidermidis* strains. This knowledge is essential to differentiate strains that cause infection from contaminating strains to shed new light on the *S. epidermidis* pathogenesis and that can be useful as a diagnostic marker.

1.5 THESIS OUTLINE

Following this introduction to this thesis, Chapter 2 provides an updated literature review with key information about *S. epidermidis* infections and their consequences on public health.

Chapters 3, 4, and 5 are the experimental chapters and address the three specific aims of this thesis. Each of the experimental chapters stands alone, providing a summary, brief introduction, materials and methods, results, and discussion.

Chapter 3 presents the optimization of a human blood *ex vivo* model to analyze *S. epidermidis* gene expression. Several parameters such as the volume of blood, the viability of leukocytes over time and concentration of bacteria are assessed in this chapter.

Chapter 4 is focused on a comparative RNA sequencing data of clinical and commensal isolates to select potential molecular markers for the diagnosis of *S. epidermidis* infections.

Chapter 5 addresses the analysis of the transcription levels of the selected potential molecular markers in a worldwide collection of *S. epidermidis* strains. The ability of these strains to survive in human blood is also assessed.

Chapter 6 summarizes the major findings, limitations of the study and future directions.

A schematic diagram that shows the outline of this thesis and the relationship of all the chapters is presented below in Figure 1.1.
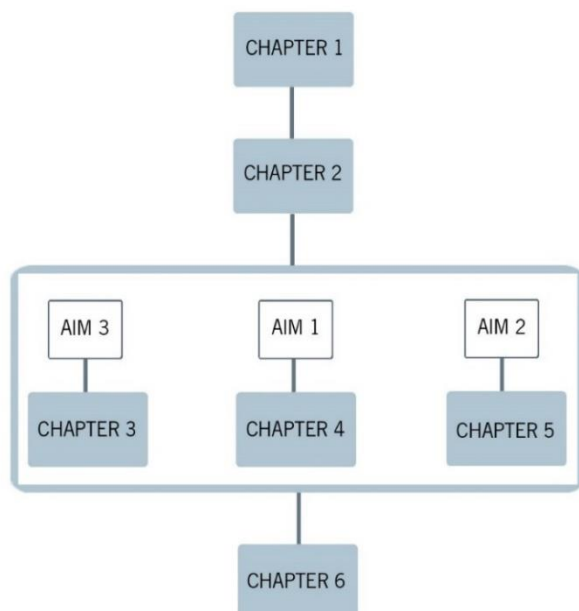
Figure 1.1. Thesis outline

1.6 REFERENCES

1. Vuong C, Otto M. *Staphylococcus epidermidis* infections. *Microbes and Infection* 2002;4:481-489.

2. Piette A, Verschraegen G. Role of coagulase-negative staphylococci in human disease. *Veterinary Microbiology* 2009;134:45-54.

3. Becker K, Heilmann C, Peters G. Coagulase-Negative Staphylococci. *Clinical Microbiology Reviews* 2014;27:870-926.

4. Hogan S, Stevens NT, Humphreys H, O'Gara JP, O'Neill E. Current and Future Approaches to the Prevention and Treatment of Staphylococcal Medical Device-Related Infections. *Current Pharmaceutical Design* 2014;21:100-113.

5. Hall KK, Lyman JA. Updated Review of Blood Culture Contamination. *Clinical Microbiology Reviews* 2006;19:788-802.

6. Alahmadi YM, Aldeyab MA, McElnay JC, Scott MG, Darwish Elhajji FW *et al.* Clinical and economic impact of contaminated blood cultures within the hospital setting. *The Journal of hospital infection* 2011;77(3):233-236.

CHAPTER 2

Literature review

SUMMARY

This chapter reviews the clinical relevance of *S. epidermidis* infections associated with the use of medical devices.

2.1 *STAPHYLOCOCCUS EPIDERMIDIS*

The genus *Staphylococcus* consists of 47 species and 23 sub-species and these can be classified based on their ability to produce the enzyme coagulase [1]. *Staphylococcus epidermidis* belongs to the coagulase-negative staphylococci (CoNS) and is one of the most prevalent species of bacteria that reside on the human skin and mucous membranes [2]. *S. epidermidis* has the ability to adapt to different characteristics of the human skin such as extended pH range, temperature, nutrients, high-salt environment and moisture content [3] and plays an important role in the maintenance of healthy skin flora [4] by interfering with the colonization by pathogenic microorganisms, such as *S. aureus* [5, 6]. However, despite being a commensal of the skin, *S. epidermidis* can also be an opportunistic pathogen in host with predisposing factors, such as preterm neonates [7], immunosuppression due to cancer treatment or HIV/AIDS [8] and the presence of indwelling medical devices [9]. In recent years, *S. epidermidis* has become the most frequent and important cause of nosocomial infections associated with the use of medical devices [10]. Moreover, *S. epidermidis* often can disseminate into the bloodstream and lead to bacteremia [11]. In fact, *S. epidermidis* is currently recognized as the most frequent cause of catheter-related bloodstream infections [12]. Furthermore, the increasing antibiotic resistance observed in *S. epidermidis* strains aggravate this problem and pose a great challenge for the management of the infections caused by this bacterium[1]. Considering that *S. epidermidis* can live at the edge between commensalism and pathogenicity, infections caused by this bacterium are difficult to diagnose [13]. For the above mentioned reasons, these infections are associated with significant healthcare costs, morbidity, and mortality [14].

2.1.1 Biofilm formation

The implantation of medical devices provides the opportunity for this bacterium to colonize and invade the host. The success of *S. epidermidis* as a pathogen is due to its ability to adhere to surfaces of medical devices [12] and form biofilms that provide resistance against host defenses [15] and antimicrobial treatment [16]. Hence, the physical removal of the colonized devices and reinsertion becomes necessary [17].

Biofilms can be defined as communities of microorganisms that adhere to each other and/or to a surface that is embedded in a self- produced extracellular matrix [18]. Biofilm formation is a multistep process and its lifecycle (depicted in Figure 2.1) involves three main stages: (1) adhesion or attachment, (2) accumulation and maturation, and (3) disassembly [19].

Adhesion   **Maturation**   **Disassembly**
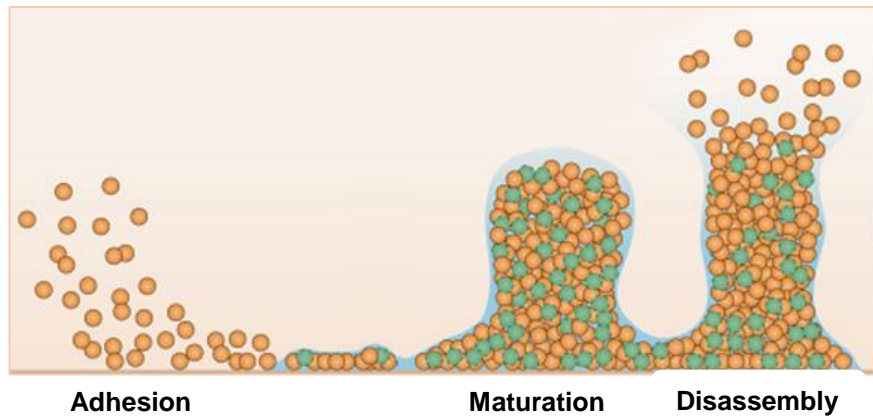
Figure 2.1. Phases of biofilm formation. Adapted from Otto [14].

2.1.1.1 Adhesion

The adhesion to the surface is the first and critical step in the development of biofilms. The adhesion can occur directly to the surface of the medical device and this process is dependent on the physicochemical properties of the bacterial surface and the solid surface [20]. Thus, this type of attachment is dependent on many physiochemical variables, such as the electrostatic and hydrophobic interactions, van der Walls forces and charge [20]. There are, however, some bacterial cell surface molecules that can play an important role in the initial adhesion such as AtlE [21] and Aae [22]. These molecules, also called microbial surface components recognizing adhesive matrix molecules, interact with host extracellular matrix components [19]. These adhesins play an important role in colonization and infection since during implantation of medical-devices the surfaces are rapidly coated by host serum proteins (e.g. fibronectin, fibrinogen, vitronectin, elastin, and collagens or host cells, such as platelets) [12]. The affinity of adhesins to such host factors may be crucial for the successful binding of staphylococci cells.

2.1.1.2 Accumulation and maturation

After initial adhesion, bacteria multiply and accumulate in a multilayered biofilm architecture [23], which is characterized by the production of an extracellular matrix, composed by polysaccharides, proteins, lipids and nucleic acids [14] and that is crucial for the stabilization of the biofilm structure [18]. The most important adhesive molecule for *S. epidermidis* biofilm formation is the polysaccharide intercellular adhesin (also known as poly-N-acetyl glucosamine; PIA/PNAG), that surrounds and connects cells in a biofilm [24]. This polysaccharide is synthesized by proteins encoded in the *ica* operon [25]. The *ica* operon is composed of four open reading frames: *icaA, icaD, icaB* and *icaC* [26]. IcaA and IcaD produce PIA/PNAG from monomers of N-acetylglucosamine [27]; followed by elongation and externalization of the full-length PIA/PNAG

molecule by IcaC protein [28]. After this, the cell-surface enzyme, IcaB, partially de-acetylates the N-acetylglucosamine residues. This process confers the cationic character that is essential to the attachment to the bacterial surface [29]. Although PIA/PNAG is an important factor in biofilm formation, *S. epidermidis* strains that lack the *ica* operon and are still able to produce biofilms and have been isolated from clinically relevant infections [30, 31]. Since then other factors that mediate biofilm formation have been identified, such as the accumulation-associated protein (Aap) [32], the homologue to the biofilm-associated protein (Bap) of *S. aureus* protein (Bhp) [33], the extracellular matrix biofilm protein (Embp) [34] and the small basic protein (Sbp) [35]. The Bhp is present in some *S. epidermidis* strains and promotes biofilm formation in the absence of PIA [36] whilst Embp is involved in adhesion and mediates biofilm accumulation [34] and Sbp is involved in intracellular adhesion and the establishment of multilayered biofilms [35]. Additionally, teichoic acids [37], a component of the surface of Gram-positive bacteria, as well as extracellular DNA (eDNA), which is released from lysed bacteria, are important adhesive molecules that allow the interaction with other molecules due to their negative charge [38]. Furthermore, disruptive forces are needed for the formation of channels through which nutrients, such as water and ions can penetrate deeper biofilms layers [37] and facilitate the removal of metabolic waste [38].

2.1.1.3 Disassembly

The final stage of the biofilm lifecycle is biofilm disassembly, also termed dispersion or detachment. The detachment of cells from the biofilms formed on medical devices is crucial for the colonization of other locations resulting, as well, in the onset of severe acute infections such as endocarditis [39], bacteremia [40] and sepsis [41]. The process of biofilm disassembly involves a combination of a complex, multi-factorial and highly regulated process that can be activated by several external signals or bacterial signals [42]. Under some specific conditions, bacteria might shift again to the planktonic lifestyle [43]. These conditions include nutrients unavailability, oxygen depletion, low levels of nitric oxide, changes of temperature, a variation of the levels of iron, accumulation of wastes [42]. However, biofilm disassembly also occur due to the passive mechanism that is mediated by external forces such as shear forces or abrasion [43]. To disperse, bacteria encased in a biofilm must be able to degrade components of the matrix in order to be released into the surrounding environment. To do so, bacteria produce a number of effector molecules, such as proteases, polysaccharide degrading enzymes, nucleases [42, 43] and phenol soluble modulins (PSMs) [44]. *S. epidermidis* produces α-type PSMs (PSMα, PSMε, PSM-mec and PSMγ (also known as δ-toxin) and β-type PSMs, which are associated with both biofilm structuring and

detachment and cytotoxicity, respectively [45]. These peptides are controlled by quorum sensing Agr system, a quorum-sensing that controls gene expression according to the increase of cell density [40, 46].

2.1.2 Antimicrobial resistance

Antibiotic resistance is a serious threat to global health since the misuse of antibiotics and the growing number of infections leads to an increase in hospital stays, costs, and mortality [47]. Resistance to different antibiotic classes is a common feature among *S. epidermidis* strains, which is often attributable to the presence of mobile genetic elements and the exposure to high antibiotic selective pressure. Antimicrobial resistance can also be acquired through horizontal gene acquisition and genetic recombination [48]. Resistance to methicillin in *S. epidermidis* is generally conferred by the presence of *mecA* gene, which encodes a penicillin-binding protein with a reduced affinity to β-lactam antibiotics [49]. *mecA* is part of a mobile genetic element designated SCC*mec* (staphylococcal cassette chromosome *mec*) [50]. Nowadays, approximately 80% of *S. epidermidis* strains isolated from device-associated infections are resistant to methicillin [14]. Healthcare workers and patients play an important role in the dissemination of methicillin-resistant *S. epidermidis* [51, 52]. Thereby, the commensal lifestyle of *S. epidermidis* renders this bacterium a reservoir for antibiotic resistance genes [53]. The acquisition of certain insertion sequence (IS) elements into the genome can also promote antimicrobial resistance. *IS256* is part of the composite transposon Tn4001, which confers aminoglycoside resistance in staphylococci [54]. Thus, the insertion sequence of *IS256* in the *S. epidermidis* chromosome has been associated with DNA rearrangements causing phenotypic changes [55]. The presence of this sequence has been associated with gentamicin and oxacillin resistance [56]. Nowadays, *S. epidermidis* strains have also acquired resistance to several other antibiotics such as rifampicin, fluoroquinolones, tetracycline, chloramphenicol, erythromycin, and sulfonamides [14].

Also contributing to the failure of antimicrobial therapy against *S. epidermidis* is its increasing antimicrobial tolerance associated with biofilm formation [57]. Interestingly, it was shown that biofilm cells maintained enhanced tolerance to different antibiotics several hours after being released from the biofilms [58]. Besides that, biofilms are very heterogeneous populations of cells with different metabolic states[59]. One of the subpopulations of biofilm cells are dormant cells that have a reduced metabolic rate [60] which enables them to evade antibiotics targeting cell wall

proteins and nucleic acids biosynthesis [61]. Biofilms have been shown to contain persisters cells (a variant of dormant cells) that are highly tolerant to antibiotics [61, 62].

2.1.3 Overview of the innate immune response during *S. epidermidis* infection

The innate immune system is the first line of defense against invading microorganisms and is composed by soluble and cellular mediators, being polymorphonuclear neutrophils (PMNs) the principal effector cells [63]. Several studies suggest that *S. epidermidis* biofilms induce attenuation in phagocytic killing by human macrophages [64] and PMNs [29, 65], as well as antibody-mediated killing by leukocytes [66].

Phagocytes activation can occur through the recognition of pathogen-associated molecular patterns on the bacterial surface by cellular recognition receptors such as Toll-like receptors (TLRs) [67]. TLRs recognize lipoproteins, lipoteichoic acid (LTA) and peptidoglycan of the cell wall [68]. *S. epidermidis* triggers immune responses via TLR2, which often forms heterodimers with TLR1 and 6 [68], which recognized staphylococci cell-wall associated molecules. Furthermore, TLR2 seems to play an important role in the clearance of *S. epidermidis* bacteremia [69]. The activation of phagocytes can also occur through the complement system [70]. Complement activation promotes the opsonization of bacteria and the recruitment of phagocyte cells to the infection site [63]. After bacterial recognition, the phagocytes internalize opsonized bacteria into the phagosome and bacteria will be destroyed by the action of reactive oxygen species (ROS) as well as antimicrobial peptides (AMPs) [71]. These peptides promote the direct killing of pathogens and cell recruitment [71].

S. epidermidis infections have a low inflammatory profile, which contributes to its chronic nature, escaping killing by the immune system and, thus, persists on the surfaces of medical devices. However, *S. epidermidis* sepsis is acute by nature and the host immune response, which takes place largely in the bloodstream, is quite different from that against *S. epidermidis* biofilm associated-infections. PIA has been described as an important biofilm constituent that promotes immune system evasion by inhibiting phagocytes killing [65]. Therefore, the presence of PIA on cell surface protects *S. epidermidis* from opsonization and PMN mediated killing [65] as well as protection against the action of AMPs [29]. However, one recent study has reported the opposite, whereby PIA producing bacteria induced pro-inflammatory response, enhanced phagocytosis [72]. Like PIA, Aap and Embp have a similar action in the protection of *S. epidermidis* against phagocytosis [73]. Additionally, *S. epidermidis* secretes proteases that have an important role in

the inactivation of host defense. The protease Esp degrades the protein C5 that provides the ability of *S. epidermidis* to evade the complement defense [6].The extracellular protease SepA can cleave and inactivate AMPs and, therefore, also promotes *S. epidermidis* protection against phagocytosis by PMNs [74]. However, the specific role of these proteases *in vivo* is difficult to define [75].

2.2 MEDICAL DEVICES-ASSOCIATED INFECTIONS

As discussed above, *S. epidermidis* is the most common pathogen in the context of nosocomial infections associated with the use of medical devices [53]. Therefore, the increased use of indwelling medical devices for diagnostic and therapeutic procedures has resulted in increment of the number of infections caused by *S. epidermidis* [76]. *S. epidermidis* infections require access through the epidermal barrier and this breach is facilitated by the implantation of intravascular devices [77]. Therefore, a medical device acts as a bridge between the external environment and the internal environment of the patient (Figure 2.2). In most cases, the contamination of the device occurs from the patient´s own skin microbiota or from the hands of the surgical or clinical staff [78]. However, there are other less frequent sources of contamination, such as catheter hub and infusate (Figure 2.2) [79].
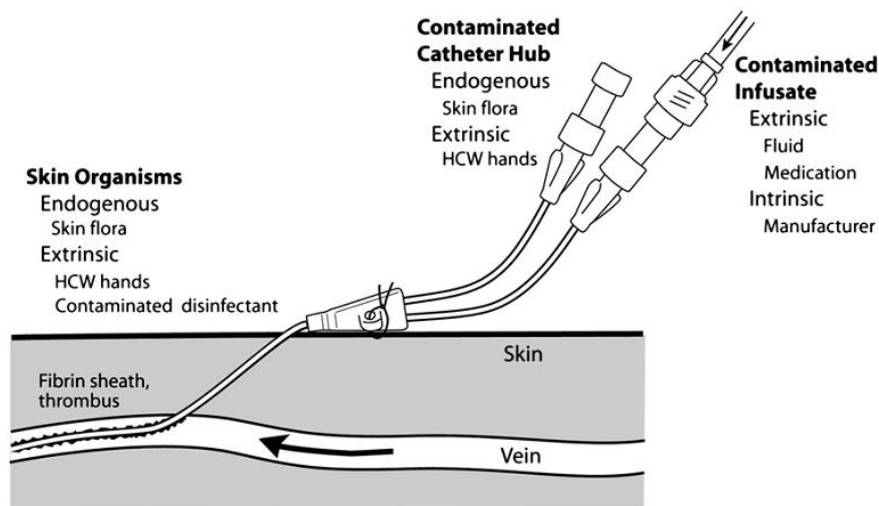
Figure 2.2. Model of infections associated with intravascular medical devices and the potential sources of contamination. HCW, healthcare worker. Adapted from Crnich et al. [79].

S. epidermidis may cause a wide range of local and systemic infections. (Table 2.1). Local infections are considered those where the medical device does not have a direct connection to the bloodstream and causes local inflammation signs such as erythema, swelling, tenderness and purulent drainage. On the other hand, systemic infections have no obvious infection focus and may

occur through bacterial dissemination either from the primary site of infection or direct entry into the bloodstream via medical devices such as catheters. Depending on nature and localization, they may cause bacteremia, sepsis, endocarditis, meningitis [1].

Table 2.1. Clinical infections caused by *S. epidermidis* and other CoNS [13]

Infections	% of CoNS
Bacteremia and intravascular catheter-related bloodstream infections	30 - 40
Endocarditis and infection of cardiac devices and vascular grafts	15 - 40
Orthopedic prosthetic device infections	30 - 43
Cerebrospinal fluid shunt infections	5
Infections of genitourinary prostheses	35 - 60
Peritoneal dialysis catheter associated infections	20 - 40

Bacteremia ranks among the seven most frequent causes of death in North America and Europe [80] and *S. epidermidis* is among the most common causative agents [10, 13]. The most frequent types of these medical device-associated infections are catheter-related bloodstream infections [1]. Immunocompromised patients and preterm neonates are primarily vulnerable to such infections [8]. since they have an immature immune system characterized by deficiencies in the complement system [81] immature neutrophils [82] and mucosal barriers with higher permeability [1]. Consequently, *S. epidermidis* predominate among the CoNS causing neonatal bacteremia [7].

2.2.1 Diagnosis of bloodstream infections

Bloodstream infections represent a growing public health concern since are associated with high rates of morbidity and mortality [12]. Diagnosis of bloodstream infections remains one of the most important challenges in the field of Medical Microbiology. Fast and accurate diagnosis is the key to avoid inappropriate therapy as well as the delay in the administration of the indicated treatment [83]. Although a variety of novel approaches have been developed over the last years to reduce the time of the identification of the microorganisms in blood cultures [83] and to increase the sensitivity and specificity [84], hemocultures remain the gold standard for the diagnosis of bloodstream infections. Successful diagnosis of bloodstream infections depends on best practices of blood culture sampling, which includes skin preparation, blood sampling site, and volume of blood, [85]. Therefore, ideally, blood samples should be obtained from peripheral venipuncture and not via vascular catheters, since blood cultures are less contaminated when samples are obtained by

peripheral venipuncture [86]. Besides, the volume of blood collected is crucial for the detection of microorganisms, since the bacterial concentration is very low in most patients with bloodstream infection [84]. An adequate volume of blood can be obtained by multiple or by single-sampling. Multi-sampling consists in collect two/three sets of blood cultures bottle (one set contain one aerobic and one anaerobic bottle) and should be sampled within 24 h [85]. On the other hand, single sampling collects the total volume of blood, through one single puncture. This strategy is preferred as it reduces the contamination ratio and improves comfort for patients [85], Then, the bottles are incubated into a continuous monitoring automated device at 35-37°C, for 5-7 days, and bacterial growth detected through the production of CO_2 that triggers a pH increase [85]. After a positive blood culture, a Gram stain is performed directly from the blood culture and, then, the pathogen is identified by biochemical tests [84].

2.2.2 The dilemma of the clinical significance: true bacteremia versus contamination of blood cultures

Due to its commensal and ubiquitous nature, CoNS contamination may occur during the blood collection process. The interpretation of CoNS positive blood cultures is complicated as it may indicate a clinically significant infection or a false-positive result without any clinical consequence. *S. epidermidis* represent the most frequent species among the CoNS isolated from blood samples, thereby, is one of the species often implicated in contaminated blood cultures and true cases of bacteremia [1, 87]. The presence of *S. epidermidis* in blood cultures presents a dilemma that can be challenging for clinicians [12]. In fact, *S. epidermidis* causes relatively indolent bloodstream infections which makes difficult to differentiate between true bacteremia from contamination [51]. Moreover, there is no clear understanding of the determinants of pathogenicity and therefore they are often described as accidental pathogen [14].

The number of blood culture samples exhibiting growth can help to differentiate contamination from bacteremia. Determination of clinical significance of bloodstream infections follows established criteria from the Centers for Disease Control and Prevention (CDC) [88]. These criteria include: (i) isolation of a known pathogenic cultured from one or more blood cultures while the isolated microorganism is not related to a concomitant infection at another site; (ii) the presence of at least one of the following signs or symptoms, when they are not related to infection at another site: fever (> 38°C), chills or hypotension; iii) patients aged ≤ 1 year with at least one of the

following signs or symptoms, when they are not related to infection at another site: - fever (> 38°C, rectal), hypothermia (<37°C, rectal) apnea, or bradycardia.

Therefore, one of the parameters used is the number of positive blood cultures to differentiate contaminated blood cultures from bacteremia. Besides the number of positive blood cultures, the time to positivity has also been explored in order to identify the significance of positive blood cultures [89]. This parameter is associated with the bacterial concentration present in the initial blood culture. It assumes that higher densities of bacteria will grow faster and, therefore, the time to positivity will be faster than samples with lower densities of bacteria. This is related to the fact that blood cultures from bacteremia are more likely to have higher bacterial densities than blood cultures from contamination [89]. Kassis *et al* [90] postulated that less than 16h to achieve positivity is associated with bacteremia and more than 20h is associated with contamination. However, Savithri *et al* [91] showed that time to positivity up to 24h can still be used as a marker of bacteremia. Therefore, regarding the, there is no consensus on using the time to positivity to evaluate positive blood cultures. In conclusion, with a combination of other parameters, the number of positive blood cultures and time to positivity can help to identify true bacteremia from contamination.

2.2.2.1 Virulence markers

In the last years, several studies have attempted to find virulence factors that could allow discrimination between isolates that cause infection and commensal isolates. The best studied virulence factor of *S. epidermidis* is its ability to form biofilms, which promotes the evasion from the host immune system and enhanced tolerance against antibiotics [92]. Therefore, the ability of clinical and commensal isolates to form biofilm has been extensively explored [93-95]. However, biofilm formation does not appear to be a suitable marker since biofilm formation can be influenced by environmental factors [96, 97], and the biofilm phenotype is subject to phase variation [93]. In addition, there is evidence that showed that clinical and commensal isolates have similar ability to form biofilm [98]. Besides differentiation through phenotypic analysis, the presence of molecular markers has been used in an attempt to identify infection-associated (Table 2.2). Many studies have used clinical isolates from different sources such as bloodstream and/or catheters-related infections, commensal isolates from healthcare workers or hospitalized patients, and commensal isolates from the community.

Several genetic determinants have been studied, regarding their potential as discriminating markers, such as the *IS256*, the *mecA* gene, and genes associated with biofilm formation associated genes (*atle, aap, bhp, fbe, icaABDC, embp, sesI, sesC, sescB, agr, luxS and sarZ*). The arginine catabolic mobile element (ACME), which increases the ability of some methicillin- resistant *Staphylococcus aureus* (MRSA) strains to grow and survive in the host [99], and the presence of the formate dehydrogenase gene (*fdh*) have also been studied. In addition, the sequencing of seven conserved housekeeping genes has been performed using multilocus sequence typing (MLST) [48, 51, 100]. In addition, the assignment of STs to genetic clusters (GC) using the Bayesian model-based clustering was used to define *S. epidermidis* population structure, which reflects the bacterial lifestyle [101, 102]. Furthermore, whole-genome sequencing (WGS), which provides a complete genome of organisms, was also used to find virulence markers that could be able to differentiate invasive from commensal isolates [103, 104]. All these studies have attempted to identify virulence factors that could be used as diagnostic markers able to differentiate between infection and false-positives blood cultures. However, none were able to achieve that goal. This is due to the fact isolates from different origins reveal an enormous genetic diversity probably due to the necessity to adapt to different environments.

Thus, the dilemma of the clinical significance of blood cultures positive for *S. epidermidis* remains. Since the ability of *S. epidermidis* to cause infection depends on its ability to survive the host microenvironment, the analysis of the differential expression of certain genes under specific conditions, such as in incubation in human blood, could be more suitable to distinguish differences between clinical and commensal isolates. This will be further discussed in Chapters 4 and 5.

Table 2.2 Summary of some of the studies performed to evaluate the potential of genetic determinants to discriminate clinical from commensal isolates

Genetic determinants/ technique	Sample	Findings	Ref
IS256	117 clinical isolates; 22 commensal isolates from medical staff; 55 commensal isolates from community.	The insertion of *IS256* was more prevalent in nosocomial isolates	[105]
atle, aap, icaA	29 commensal isolates from healthy volunteers; 16 clinical isolates from catheters and 36 from blood culture	The presence of *aap* and *icaA* was more prevalent in clinical isolates from catheters but cannot be used for clinical decision	[106]
ica, mecA, IS256	41 clinical isolates from blood cultures; 42 commensal isolates from patients; 15 commensal isolates from healthy volunteers	The presence of these genes does not allow proper discrimination between invasive from colonization	[107]
icaA, aap	67 clinical isolates from blood cultures and 38 from catheters; 60 commensal isolates from healthy individuals	Minor difference between isolates from different origins	[108]
fdh, mecA	28 clinical isolates from catheters and blood cultures; 71 commensal isolates from healthy individuals	*fdh* gene was more prevalent in commensal isolates and has been proposed as a discriminatory marker	[103]
aap, bhp, atle, fbe	19 clinical isolates from infection; 38 contaminants isolates	The presence of bap gene seems to be associated with colonization	[109]

Table 2.2. Continued

Genetic determinants/ technique	Sample	Findings	Ref.
hld, agr, mecA, aap, icaADB	61 clinical isolates; 24 commensal isolates from healthy individuals	No differences in the detection of these genes	[111]
ica, bhp, IS256, altE, fbe, embp, aap	47 isolates from catheter-related infection; 25 isolates form bacteremia; 25 commensal isolates from catheters hubs	*ica* gene and *IS256* were more prevalent in isolates that cause infection. *aap, fbe, altE, embp* and *bhp* genes were equally distributed in invasive and commensal strains	[95]
sesI	54 isolates from infection; 29 isolates from contamination; 24 commensal isolates from healthy individuals	The presence of this gene cannot be used since it was only found in 50% of clinical isolates	[112]
ACME	33 clinical isolates from bloodstream infections; 33 commensal isolates from healthy individuals	ACME was more prevalent in commensal isolates	[113]
ACME, mecA, 13 biofilm- associated genes	50 isolates from bloodstream infection; 19 commensal isolates from nasal	sesI and sdrf were more prevalent in isolates from bloodstream infections	[114]
sesB, sesC, sesI, aap, atlE, embp, icaA, icaC, mecA, ACME, IS256, sarZ	80 isolates from bacteremia; 80 isolates from contamination	The presence of this virulence determinants cannot be used to distinguish bacteremia from contamination	[104]

Table 2.2. Continued

Genetic determinants/ technique	Sample	Findings	Ref.
luxS	43 clinical isolates from blood and 6 from catheters; 35 commensal isolates from healthy individuals	No significant differences were found in the expression of luxS gene between clinical and commensal	[110]
MLST	50 clinical isolates from true bacteremia; 25 contaminants isolates; 25 commensal isolates	ST2 and ST5 were more prevalent in isolates from true bacteremia and contaminants	[100]
MLST	20 clinical isolates from blood cultures; 42 commensal isolates from healthcare workers	ST2 and ST4 were more prevalent in clinical isolates	[51]
Genetic Clusters	59 clinical isolates from bacteremia; 55 isolates from blood culture contamination; 40 isolates from non-hospital subjects	GC predicted isolates from non-hospital sources from hospital sources in 80%, but not discriminate infection from contamination	[102]
WGS	80 isolates from bacteremia; 80 isolates from contamination	Only the presence of IS256 was prevalent in isolates from bacteremia	[104]

2.3 REFERENCES

1. **Becker K, Heilmann C, Peters G**. Coagulase-Negative Staphylococci. *Clinical Microbiology Reviews* 2014;27:870-926.

2. **Vuong C, Otto M**. *Staphylococcus epidermidis* infections. *Microbes and Infection* 2002;4:481-489.

3. **Rogers LK, Fey PD, Rupp ME**. Epidemiology of Coagulase-Negative Staphylococci and Infections Caused by These Organisms. *Staphylococci in Human Disease*: Blackwell Publishing Ltd; 2009. pp. 310-332.

4. **Otto M**. Molecular basis of *Staphylococcus epidermidis* infections. *Seminars in Immunopathology* 2012;34:201-214.

5. **Iwase T, Uehara Y, Shinji H, Tajima A, Seo H** *et al.* *Staphylococcus epidermidis* Esp inhibits *Staphylococcus aureus* biofilm formation and nasal colonization. *Nature* 2010;465(7296):346-349.

6. **Sugimoto S, Iwamoto T, Takada K, Okuda K, Tajima A** *et al.* *Staphylococcus epidermidis* Esp degrades specific proteins associated with *Staphylococcus aureus* biofilm formation and host-pathogen interaction. *Journal of Bacteriology* 2013;195(8):1645-1655.

7. **Dong Y, Speer CP**. The role of *Staphylococcus epidermidis* in neonatal sepsis: Guarding angel or pathogenic devil? *International Journal of Medical Microbiology* 2014;304(5):513-520.

8. **Nguyen TH, Park MD, Otto M**. Host Response to *Staphylococcus epidermidis* Colonization and Infections. *Front Cell Infect Microbiology*, Review 2017;7:90.

9. **McCann MT, Gilmore BF, Gorman SP**. *Staphylococcus epidermidis* device-related infections: pathogenesis and clinical management. *Journal of Pharmacy and Pharmacology* 2008;60:1551-1571.

10. **Kleinschmidt S, Huygens F, Faoagali J, Rathnayake IU, Hafner LM**. *Staphylococcus epidermidis* as a cause of bacteremia. *Future Microbiology* 2015;10(11):1859-1879.

11. **Wisplinghoff H, Rosato AE, Enright MC, Noto M, Craig W** *et al.* Related clones containing SCCmec type IV predominate among clinically significant *Staphylococcus epidermidis* isolates. *Antimicrob Agents Chemotherapy* 2003;47(11):3574-3579.

12. **Buttner H, Mack D, Rohde H**. Structural basis of *Staphylococcus epidermidis* biofilm formation: mechanisms and molecular interactions. *Frontiers in Cellular and Infection Microbiology* 2015;5.

13. **Rupp ME**. Clinical Characteristics of Infections in Humans Due to *Staphylococcus epidermidis*. In: Fey PD (editor). *Staphylococcus Epidermidis*. Totowa, NJ: Humana Press; 2014. pp. 1-16.

14. **Otto M**. *Staphylococcus epidermidis*–the 'accidental' pathogen. *Nature Reviews Microbiology* 2009;7(8):555-567.

15. **Cerca N, Jefferson KK, Oliveira R, Pier GB, Azeredo J**. Comparative Antibody-Mediated Phagocytosis of *Staphylococcus epidermidis* Cells Grown in a Biofilm or in the Planktonic State. *Infection and Immunity* 2006;74(8):4849-4855.

16. **Cerca N**. Comparative assessment of antibiotic susceptibility of coagulase-negative staphylococci in biofilm versus planktonic culture as assessed by bacterial enumeration or rapid XTT colorimetry. *Journal of Antimicrobial Chemotherapy* 2005;56:331-336.

17. **Mack D, Davies AP, Harris LG, Jeeves R, Pascoe B** *et al.* *Staphylococcus epidermidis* in Biomaterial-Associated Infections. In: Moriarty TF, Zaat SAJ, Busscher HJ (editors). *Biomaterials Associated Infection*. New York, NY: Springer New York; 2013. pp. 25-56.

18. **Costerton JW**. Bacterial Biofilms: A Common Cause of Persistent Infections. *Science* 1999;284:1318-1322.

19. **Fey PD, Olson ME**. Current concepts in biofilm formation of *Staphylococcus epidermidis*. *Future Microbiol* 2010;5(6):917-933.

20. **Vacheethasanee K, Temenoff JS, Higashi JM, Gary A, Anderson JM** *et al.* Bacterial surface properties of clinically isolated *Staphylococcus epidermidis* strains determine adhesion on polyethylene. *Journal of Biomedical Materials Research* 1998;42:425-432.

21. **Heilmann C, Hussain M, Peters G, Gotz F**. Evidence for autolysin-mediated primary attachment of *Staphylococcus epidermidis* to a polystyrene surface. *Molecular Microbiology* 1997;24:1013-1024.

22. **Heilmann C**. Identification and characterization of a novel autolysin (Aae) with adhesive properties from *Staphylococcus epidermidis*. *Microbiology* 2003;149:2769-2778.

23. **von Eiff C, Heilmann C, Peters G**. New aspects in the molecular basis of polymer-associated infections due to staphylococci. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology* 1999;18:843-846.

24. **Mack D, Fischer W, Krokotsch A, Leopold K, Hartmann R** *et al.* The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked

glucosaminoglycan: purification and structural analysis. *Journal of Bacteriology* 1996;178:175-183.

25. Heilmann C, Schweitzer O, Gerke C, Vanittanakom N, Mack D *et al.* Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Molecular Microbiology* 1996;20:1083-1091.

26. O'Gara JP, Humphreys H. *Staphylococcus epidermidis* biofilms: importance and implications. *Journal of Medical Microbiology* 2001;50:582-587.

27. Gerke C, Kraft A, Sussmuth R, Schweitzer O, Gotz F. Characterization of the N-acetylglucosaminyltransferase activity involved in the biosynthesis of the *Staphylococcus epidermidis* polysaccharide intercellular adhesin. *Journal of Biological Chemistry* 1998;273(29):18586-18593.

28. Rohde H, Kalitzky M, Kroger N, Scherpe S, Horstkotte MA *et al.* Detection of Virulence-Associated Genes Not Useful for Discriminating between Invasive and Commensal *Staphylococcus epidermidis* Strains from a Bone Marrow Transplant Unit. *Journal of Clinical Microbiology* 2004;42:5614-5619.

29. Vuong C, Voyich JM, Fischer ER, Braughton KR, Whitney AR *et al.* Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. *Cellular Microbiology* 2004;6:269-275.

30. Vandecasteele SJ, Peetermans WE, Merckx R, Van Eldere J. Expression of biofilm-associated genes in *Staphylococcus epidermidis* during in vitro and in vivo foreign body infections. *Journal Infection Disease* 2003;188(5):730-737.

31. Kogan G, Sadovskaya I, Chaignon P, Chokr A, Jabbouri Sd. Biofilms of clinical strains of *Staphylococcus* that do not contain polysaccharide intercellular adhesin. *FEMS Microbiology Letters* 2006;255:11-16.

32. Rohde H, Burdelski C, Bartscht K, Hussain M, Buck F *et al.* Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases: Aap mediated *S. epidermidis* biofilm formation. *Molecular Microbiology* 2005;55:1883-1895.

33. Tormo MA, Knecht E, Gotz F, Lasa I, Penades JR. Bap-dependent biofilm formation by pathogenic species of *Staphylococcus*: evidence of horizontal gene transfer? *Microbiology* 2005;151(Pt 7):2465-2475.

34. Christner M, Franke GC, Schommer NN, Wendt U, Wegert K *et al.* The giant extracellular matrix-binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment to fibronectin. *Molecular Microbiology* 2010;75:187-207.

35. Decker R, Burdelski C, Zobiak M, Büttner H, Franke G *et al.* An 18 kDa Scaffold Protein Is Critical for *Staphylococcus epidermidis* Biofilm Formation. *PLOS Pathogens* 2015;11:e1004735.

36. Lasa I, Penadés JR. Bap: A family of surface proteins involved in biofilm formation. *Research in Microbiology* 2006;157:99-107.

37. Watnick P, Kolter R. Biofilm, City of Microbes. *Journal of Bacteriology* 2000;182:2675-2679.

38. Habash M, Reid G. Microbial biofilms: their development and significance for medical device-related infections. *Journal of Clinical Pharmacology*, Review 1999;39(9):887-898.

39. Elgharably H, Hussain ST, Shrestha NK, Blackstone EH, Pettersson GB. Current Hypotheses in Cardiac Surgery: Biofilm in Infective Endocarditis. *Semin Thorac Cardiovasc Surg Spring*, Review 2016;28(1):56-59.

40. Wang R, Khan BA, Cheung GYC, Bach T-HL, Jameson-Lee M *et al. Staphylococcus epidermidis* surfactant peptides promote biofilm maturation and dissemination of biofilm-associated infection in mice. *Journal of Clinical Investigation* 2011;121:238-248.

41. Rumbaugh DFK. The Consequences of Biofilm Dispersal on the Host. *Scientific reports* 2018;10738(10738).

42. McDougald D, Rice SA, Barraud N, Steinberg PD, Kjelleberg S. Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nature Reviews Microbiology* 2011;10(1):39-50.

43. Kaplan JB. Biofilm Dispersal: Mechanisms, Clinical Implications, and Potential Therapeutic Uses. *Journal of Dental Research* 2010;89:205-218.

44. Boles BR, Horswill AR. Staphylococcal biofilm disassembly. *Trends in Microbiology* 2011;19:449-455.

45. Cheung GY, Otto M. Understanding the significance of *Staphylococcus epidermidis* bacteremia in babies and children. *Current Opinion in Infection Disease* 2010;23(3):208-216.

46. Vuong C, Durr M, Carmody AB, Peschel A, Klebanoff SJ *et al.* Regulated expression of pathogen-associated molecular pattern molecules in *Staphylococcus epidermidis*: quorum-sensing determines pro-inflammatory capacity and production of phenol-soluble modulins. *Cell Microbiol*, Comparative Study 2004;6(8):753-759.

47.   Organization WH. Antibiotic Resistance2016.

48. **Miragaia M, Thomas JC, Couto I, Enright MC, de Lencastre H**. Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing data. *Journal of Bacteriology* 2007;189(6):2540-2552.

49. **Hartman BJ, Tomasz A**. Low-affinity penicillin-binding protein associated with beta-lactam resistance in *Staphylococcus aureus*. *Journal of Bacteriology*1984;158(2):513-516.

50. **Donlan RM, Costerton JW**. Biofilms: Survival Mechanisms of Clinically Relevant Microorganisms. *Clinical Microbiology Reviews* 2002;15:167-193.

51. **Cherifi S, Byl B, Deplano A, Nagant C, Nonhoff C** *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. *Annals of Clinical Microbiology Antimicrobials* 2014;13(20):1476-0711.

52. **Widerstrom M, Wistrom J, Edebro H, Marklund E, Backman M** *et al.* Colonization of patients, healthcare workers, and the environment with healthcare-associated *Staphylococcus epidermidis* genotypes in an intensive care unit: a prospective observational cohort study. *BMC Infection Disease*, Observational Study 2016;16(1):016-2094.

53. **Otto M**. Staphylococcal Biofilms. In: Romeo T (editor). *Bacterial Biofilms*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. pp. 207-228.

54. **Lyon BR, May JW, Skurray RA**. Tn4001: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Molecular Genomics Gene*tics, 1984;193(3):554-556.

55. **Ziebuhr W, Krimmer V, Rachid S, Lößner I, Götz F** *et al.* A novel mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS256. *Molecular Microbiology* 1999;32(2):345-356.

56. **Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K** *et al.* The bacterial insertion sequence element IS256 occurs preferentially in nosocomial *Staphylococcus epidermidis* isolates: association with biofilm formation and resistance to aminoglycosides. *Infection and Immunity* 2004;72(2):1210-1215.

57. **Zheng Z, Stewart PS**. Penetration of Rifampin through *Staphylococcus epidermidis* Biofilms. *Antimicrobial Agents and Chemotherapy* 2002;46:900-903.

58. **Gaio V, Cerca N**. Cells released from *S. epidermidis* biofilms present increased antibiotic tolerance to multiple antibiotics. *PeerJ* 2019;15(7).

59. **Stewart PS**. Antimicrobial Tolerance in Biofilms. *Microbiology Spectrum*, Meta-Analysis 2015;3(3):0010-2014.

60. **Carvalhais V, França A, Cerca F, Vitorino R, Pier GB** *et al.* Dormancy within *Staphylococcus epidermidis* biofilms: a transcriptomic analysis by RNA-seq. *Applied Microbiology and Biotechnology* 2014;98:2585-2596.

61. **Cerca F, Franca A, Perez-Cabezas B, Carvalhais V, Ribeiro A** *et al.* Dormant bacteria within *Staphylococcus epidermidis* biofilms have low inflammatory properties and maintain tolerance to vancomycin and penicillin after entering planktonic growth. *Journal Medical Microbiology* 2014;63(Pt 10):1274-1283.

62. **Lewis K**. Persister cells: molecular mechanisms related to antibiotic tolerance. *Handbook Experimental Pharmacology* 2012;211:121-133.

63. **Le KY, Park MD, Otto M**. Immune Evasion Mechanisms of *Staphylococcus epidermidis* Biofilm Infection. *Frontiers Microbiology*, Review 2018;9(359).

64. **Spiliopoulou AI, Kolonitsiou F, Krevvata MI, Leontsinidis M, Wilkinson TS** *et al.* Bacterial adhesion, intracellular survival and cytokine induction upon stimulation of mononuclear cells with planktonic or biofilm phase *Staphylococcus epidermidis*. *FEMS Microbiology Letters* 2012;330(1):56-65.

65. **Kristian SA, Birkenstock TA, Sauder U, Mack D, Gotz F** *et al.* Biofilm formation induces C3a release and protects *Staphylococcus epidermidis* from IgG and complement deposition and from neutrophil-dependent killing. *Journal of Infection and Disease* 2008;197(7):1028-1035.

66. **Cerca N, Jefferson KK, Oliveira R, Pier GB, Azeredo J**. Comparative Antibody-Mediated Phagocytosis of *Staphylococcus epidermidis* Cells Grown in a Biofilm or in the Planktonic State. *Infection and Immunity* 2006;74:4849-4855.

67. **Akira S, Hemmi H**. Recognition of pathogen-associated molecular patterns by TLR family. *Immunology Letters*, Review 2003;85(2):85-95.

68. **Fournier B**. The function of TLR2 during staphylococcal diseases. *Frontiers in Cellular Infection Microbiology* 2013;2(167).

69. **Strunk T, Power Coombs MR, Currie AJ, Richmond P, Golenbock DT** *et al.* TLR2 mediates recognition of live *Staphylococcus epidermidis* and clearance of bacteremia. *PLOS ONE* 2010;5(4):0010111.

70. **Rus H, Cudrici C, Niculescu F**. The role of the complement system in innate immunity. *Immunologic Research* 2005;33(2):103-112.

71. **Sabate Bresco M, Harris LG, Thompson K, Stanic B, Morgenstern M** *et al.* Pathogenic Mechanisms and Host Interactions in *Staphylococcus epidermidis* Device-Related Infection. *Frontiers Microbiology*, Review 2017;8(1401).

72. **Ferreirinha P, Perez-Cabezas B, Correia A, Miyazawa B, Franca A** *et al.* Poly-N-Acetylglucosamine Production by *Staphylococcus epidermidis* Cells Increases Their In Vivo Proinflammatory Effect. *Infection Immunity*, 2016;84(10):2933-2943.

73. **Schommer NN, Christner M, Hentschke M, Ruckdeschel K, Aepfelbacher M** *et al.* *Staphylococcus epidermidis* Uses Distinct Mechanisms of Biofilm Formation To Interfere with Phagocytosis and Activation of Mouse Macrophage-Like Cells 774A.1. *Infection and Immunity* 2011;79:2267-2276.

74. **Cheung GYC, Rigby K, Wang R, Queck SY, Braughton KR** *et al. Staphylococcus epidermidis* Strategies to Avoid Killing by Human Neutrophils. *PLoS Pathogens* 2010;6:e1001133.

75. **Otto M**. Staphylococcal Biofilms. *Microbiology Spectrum*, Review 2018;6(4):0023-2018.

76. **Piette A, Verschraegen G**. Role of coagulase-negative staphylococci in human disease. *Veterinary Microbiology* 2009;134:45-54.

77. **Hogan S, Stevens NT, Humphreys H, O'Gara JP, O'Neill E**. Current and Future Approaches to the Prevention and Treatment of Staphylococcal Medical Device-Related Infections. *Current Pharmaceutical Design* 2014;21:100-113.

78. **von Eiff C, Jansen B, Kohnen W, Becker K**. Infections associated with medical devices: pathogenesis, management and prophylaxis. *Drugs* 2005;65(2):179-214.

79. **Crnich CJ, Drinka P**. Medical Device;Associated Infections in the Long-Term Care Setting. *Infectious Disease Clinics* 2012;26(1):143-164.

80. **Goto M, Al-Hasan MN**. Overall burden of bloodstream infection and nosocomial bloodstream infection in North America and Europe. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2013;19(6):501-509.

81. **Granslo HN, Klingenberg C, Fredheim EA, Acharya G, Mollnes TE** *et al. Staphylococcus epidermidis* biofilms induce lower complement activation in neonates as compared with adults. *Pediatr Res* 2013;73(3):294-300.

82. **Bjorkqvist M, Kallman J, Fjaertoft G, Xu S, Venge P** *et al.* Human neutrophil lipocalin: normal levels and use as a marker for invasive infection in the newborn. *Acta Paediatric*, Comparative Study 2004;93(4):534-539.

83. **Dubourg G, Raoult D**. Emerging methodologies for pathogen identification in positive blood culture testing. *Expert Review of Molecular Diagnostic* 2016;16(1):97-111.

84. **Opota O, Croxatto A, Prod'hom G, Greub G**. Blood culture-based diagnosis of bacteraemia: state of the art. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2015;21(4):313-322.

85. **Lamy B, Dargère S, Arendrup MC, Parienti JJ, Tattevin P**. How to Optimize the Use of Blood Cultures for the Diagnosis of Bloodstream Infections? A State-of-the Art. *Frontiers Microbiology*, Review 2016;7(697).

86. **Snyder SR, Favoretto AM, Baetz RA, Derzon JH, Madison BM** *et al.* Effectiveness of practices to reduce blood culture contamination: a Laboratory Medicine Best Practices systematic review and meta-analysis. *Clinical Biochemistry* 2012;45(13-14):999-1011.

87. **Bradford R, Abdul Manan R, Daley AJ, Pearce C, Ramalingam A** *et al.* Coagulase-negative staphylococci in very-low-birth-weight infants: inability of genetic markers to distinguish invasive strains from blood culture contaminants. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2006;25(5):283-290.

88. **Tokars JI**. Predictive Value of Blood Cultures Positive for Coagulase-Negative Staphylococci: Implications for Patient Care and Health Care Quality Assurance. *Clinical Infectious Diseases* 2004;39(3):333-341.

89. **Hall KK, Lyman JA**. Updated Review of Blood Culture Contamination. *Clinical Microbiology Reviews* 2006;19:788-802.

90. **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. *Journal Clinical Microbiology* 2009;47(10):3255-3260.

91. **Savithri MB, Iyer V, Jones M, Yarwood T, Looke D** *et al.* Epidemiology and significance of coagulase-negative staphylococci isolated in blood cultures from critically ill adult patients. *Critical Care Resuscitation*, Comparative Study 2011;13(2):103-107.

92. **Schommer NN, Christner M, Hentschke M, Ruckdeschel K, Aepfelbacher M** *et al.* *Staphylococcus epidermidis* uses distinct mechanisms of biofilm formation to interfere with phagocytosis and activation of mouse macrophage-like cells 774A.1. *Infection Immunity*, 2011;79(6):2267-2276.

93. Ziebuhr W, Heilmann C, Götz F, Meyer P, Wilms K *et al.* Detection of the intercellular adhesion gene cluster (ica) and phase variation in *Staphylococcus epidermidis* blood culture strains and mucosal isolates. *Infection and Immunity* 1997;65(3):890-896.

94. Galdbart J-O, Allignet J, Tung H-S, Rydèn C, El Solh N. Screening for *Staphylococcus epidermidis* Markers Discriminating between Skin-Flora Strains and Those Responsible for Infections of Joint Prostheses. *The Journal of Infectious Diseases* 2000;182(1):351-355.

95. Mekni MA, Bouchami O, Achour W, Ben Hassen A. Strong biofilm production but not adhesion virulence factors can discriminate between invasive and commensal *Staphylococcus epidermidis* strains. *APMIS* 2012;120(8):605-611.

96. Barker LP, Simpson WA, Christensen GD. Differential production of slime under aerobic and anaerobic conditions. *Journal Clinical and Microbiology* 1990;28(11):2578-2579.

97. Deighton M, Borland R. Regulation of slime production in *Staphylococcus epidermidis* by iron limitation. *Infection and Immunity* 1993;61(10):4473-4479.

98. Freitas AI, Lopes N, Oliveira F, Bras S, Franca A *et al.* Comparative analysis between biofilm formation and gene expression in *Staphylococcus epidermidis* isolates. *Future Microbiology* 2018;13:415-427.

99. Shore AC, Rossney AS, Brennan OM, Kinnevey PM, Humphreys H *et al.* Characterization of a novel arginine catabolic mobile element (ACME) and staphylococcal chromosomal cassette mec composite island with significant homology to *Staphylococcus epidermidis* ACME type II in methicillin-resistant Staphylococcus aureus genotype ST22-MRSA-IV. *Antimicrobials Agents Chemotherapy* 2011;55(5):1896-1905.

100. Sharma P, Satorius AE, Raff MR, Rivera A, Newton DW *et al.* Multilocus Sequence Typing for Interpreting Blood Isolates of *Staphylococcus epidermidis*. *Interdisciplinary Perspectives on Infectious Diseases* 2014;2014:1-4.

101. Thomas JC, Zhang L, Robinson DA. Differing lifestyles of *Staphylococcus epidermidis* as revealed through Bayesian clustering of multilocus sequence types. *Infection Genetics Evolution*, 2014;22:257-264.

102. Tolo I, Thomas JC, Fischer RSB, Brown EL, Gray BM *et al.* Do *Staphylococcus epidermidis* genetic clusters predict isolation sources? *Journal of Clinical Microbiology* 2016;54(7):1711-1719.

103. Conlan S, Mijares LA, Becker J, Blakesley RW, Bouffard GG *et al.* *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biology* 2012;13(7):2012-2013.

104. Kleinschmidt SL. *Molecular characterisation of Bacteraemia and contamination isolates of Staphylococcus epidermidis isolated from patient blood cultures.* 2018.

105. Gu J, Li H, Li M, Vuong C, Otto M *et al.* Bacterial insertion sequence IS256 as a potential molecular marker to discriminate invasive strains from commensal strains of *Staphylococcus epidermidis*. *Journal of Hospital Infection* 2005;61:342-348.

106. Vandecasteele SJ, Peetermans WE, Merckx RR, Rijnders BJA, Van Eldere J. Reliability of the *ica, aap* and *atlE* genes in the discrimination between invasive, colonizing and contaminant *Staphylococcus epidermidis* isolates in the diagnosis of catheter-related infections. *Clinical Microbiology and Infection* 2003;9(2):114-119.

107. Rohde H, Kalitzky M, Kroger N, Scherpe S, Horstkotte MA *et al.* Detection of virulence-associated genes not useful for discriminating between invasive and commensal *Staphylococcus epidermidis* strains from a bone marrow transplant unit. *Journal Clinical Microbiology* 2004;42(12):5614-5619.

108. Liduma I, Tracevska T, Bers U, Zilevica A. Phenotypic and genetic analysis of biofilm formation by *Staphylococcus epidermidis*. *Medicina* 2012;48(6):305-309.

109. Papadimitriou-Olivgeri I, Giormezis N, Papadimitriou-Olivgeris M, Zotou A, Kolonitsiou F *et al.* Number of positive blood cultures, biofilm formation, and adhesin genes in differentiating true coagulase-negative staphylococci bacteremia from contamination. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2016;35(1):57-66.

110. Tracevska T, Liduma I, Bers U, Riekstina U, Zilevica A. Evaluation of the expression of LuxS gene in clinical isolates of *Staphylococcus epidermidis* from bloodstream infections. *Innovative Journal of Medical and Health Science* 2012:59-62.

111. Hellmark B, Soderquist B, Unemo M, Nilsdotter-Augustinsson A. Comparison of *Staphylococcus epidermidis* isolated from prosthetic joint infections and commensal isolates in regard to antibiotic susceptibility, agr type, biofilm production, and epidemiology. *International Journal of Medical Microbiology* 2013;303(1):32-39.

112. Soderquist B, Andersson M, Nilsson M, Nilsdotter-Augustinsson A, Persson L *et al.* *Staphylococcus epidermidis surface protein I (SesI): a marker of the invasive capacity of S. epidermidis?*. J Med Microbiol. 2009;58:1395-7.

113. Cherifi S, Byl B, Deplano A, Nonhoff C, Denis O *et al.* Comparative Epidemiology of *Staphylococcus epidermidis* Isolates from Patients with Catheter-Related Bacteremia and from Healthy Volunteers. *Journal of Clinical Microbiology* 2013;51:1541-1547.

114. Salgueiro VC, Iorio NL, Ferreira MC, Chamon RC, Dos Santos KR. Methicillin resistance and virulence genes in invasive and nasal *Staphylococcus epidermidis* isolates from neonates. *BMC Microbiology* 2017;17(1):017-0930

CHAPTER 3

Optimizing a reliable *ex vivo* human blood model to analyze *S. epidermidis* gene expression

SUMMARY

Human blood is often used as an *ex vivo* model to mimic the environment encountered by pathogens inside the host. A significant variety of experimental set-up conditions have been reported however, optimization strategies are often not described. This chapter aimed to evaluate parameters that are expected to influence *S. epidermidis* gene expression when using human blood *ex vivo* models. The data obtained demonstrated that a loss of 2% in leukocytes viability after 4h of blood collection resulted in a 5-fold loss of antimicrobial activity against *S. epidermidis*. In higher bacterial inocula this had a significant effect. Furthermore, it was demonstrated that the volume of human blood could be reduced to as little as 0.18 mL without affecting neither the stability of gene transcription nor the percentage of bacterial survival. Overall, the data described herein highlight experimental parameters that should be considered during *ex vivo* gene expression studies.

3.1 BRIEF INTRODUCTION

As described in the previous chapter, *S. epidermidis* can originate medical device-associated bloodstream infections [1]. Due to the clinical relevance of these infections, it is urgent to understand the strategies employed by *S. epidermidis* to endure the host immune system response. In order to better comprehend how *S. epidermidis* adapts to the host, human blood has been used as an *ex vivo* model in an attempt to mimic the *in vivo* environment [2] and to study *S. epidermidis* behavior [3, 4]. This model has also been used to study several other clinically relevant microorganisms, including *Staphylococcus aureus* [5], *Neisseria meningitidis* [6, 7], *Streptococcus agalactiae* [8], *Enterococcus faecallis* [9] and *Candida albicans* [10-12]. Human blood *ex vivo* models have contributed to a better understanding of how pathogens survive in human blood by evaluating the transcriptional response during incubation in human blood, as well as by exploring host-pathogen interactions.

To pursue the main research question of this thesis, a large number of experiments using human blood will be required. Therefore, it is important to minimize the amount of human blood used in each assay, while maintaining stable experimental results. Human blood *ex vivo* model is relatively easy to implement. However, often, the implementation of this model lack optimization steps and, not surprisingly, a significant variety of experimental set-up conditions have been reported [3, 4, 10, 11]. Because different experimental designs can greatly influence experimental conclusions, we became interested in evaluating parameters that can compromise *S. epidermidis* gene expression studies when using a human blood *ex vivo* model. To achieve this goal, the influence of the bacterial concentration in co-incubation and the effect of the viability of blood leukocytes on the ability of bacteria to survive in human blood were evaluated. Besides that, it was also investigated if reducing the volume of human blood used in *ex vivo* models would have an impact on the stability of *S. epidermidis* gene expression.

3.2 MATERIAL AND METHODS

3.2.1 Bacterial strains and growth conditions

S. epidermidis PT12003, isolated from a patient with a central catheter after stomach surgery [13], was used in this study. One single colony was inoculated into 2 mL of Tryptic Soy Broth (TSB) (Liofilchem, Teramo, Italy) and incubated overnight at 37∘C and at 120 rpm (ES-20 Shaker-Incubator, Biosan, Riga, Latvia). Planktonic cultures were started by adjusting the optical density

(OD), at 640 nm, to 0.05 ($\pm$ 0.005) in 10 mL of TSB and grown, in a 25 mL flask, for 6h at 37°C and at 120 rpm. The suspension was then washed once and resuspended in 0.9% NaCl in order to obtain an inoculum of 1×10^9 CFU/mL.

3.2.2 Human blood collection

Peripheral blood was collected from healthy adult volunteers (7 female and 4 male donors) not taking antibiotics or anti-inflammatory medication within the last 14 days, by venipuncture. Blood was drawn using one of the following anticoagulants tubes: K$_3$EDTA, sodium citrate (Vacuette, Greiner Bio-one, Kremsmünster, Austria) or lithium heparin (Becton Dickinson, NJ, USA). Blood was collected under a protocol approved by the Institutional Review Board of the University of Minho (SECVS 002/2014 (ADENDA)), which is in strict accordance with the Declaration of Helsinki and Oviedo Convention. All donors gave written informed consent to have blood taken.

3.2.3 The influence of anticoagulants on bacterial growth

In order to explore the effect of different anticoagulants on bacterial growth, TSB (Liofilchem, Teramo, Italy) was added to the different blood collection tubes and were shaken for the same duration as the tubes containing the human blood, before being transferred to a 25 mL flask. Further, TSB was transferred to Erlenmeyer flasks and bacteria were then inoculated into TSB with the different anticoagulants, obtaining a final concentration of 10^7 CFU/mL. Negative control was accomplished by inoculating bacteria in TSB without anticoagulants. These suspensions were incubated for 24h at 37°C and 120 rpm (ES-20 Shaker-Incubator). Bacterial growth was determined by CFU quantification by taking aliquots every 2h.

3.2.4 The ability of bacteria to survive in human blood

For the evaluation of the ability of bacteria to survive in human blood, 0.1 mL of different concentrations of exponentially growing bacteria, were added to 0.9 mL of human blood in 2 mL tubes, to obtain concentrations of 10^8, 10^7, 10^6, 10^5 and 10^4 colony-forming units per mL (CFU/mL) and incubated at 37°C, at 80 rpm, for up to 8h. Four and 8h after incubation, the enumeration of bacteria was determined by CFU counting. The number of CFU/mL immediately (i.e., less than 2 min) after incubation with blood was used as a control to calculate the percentage of survival. This experiment was performed five independent times, using blood from five different donors.

3.2.5 Viability of human blood leukocytes overtime

For the analysis of human blood leukocytes viability overtime (in the absence of bacteria), whole blood was collected and incubated for up to 8h at 37°C and 80 rpm (PSU-10i, Biosan). At time

points 4 and 8h, 2 mL of whole blood were collected and incubated with 5 mL of red blood cells (RBC) lysis buffer (Alfa Aesar, Karlsruhe, Germany). The suspension was mixed by carefully inverting the tubes and then incubated at room temperature during 10 min. The reaction was stopped by adding 15 mL of phosphate buffered saline (PBS) (Gibco, MA, USA). Leukocytes were harvested by 10 min centrifugation at 300 g and 4°C, and a new RBC lysis cycle was performed to lyse residual red blood cells. Leukocytes were then suspended in 0.5 mL of PBS and cells viability determined through flow cytometry (EC800, Sony Biotechnologies Inc, CA, USA), using propidium iodide staining (5 µg/mL, Sigma, MO, USA). The same procedure was performed using blood immediately after collection. This experiment was performed three independent times, using blood from different donors.

3.2.6 The impact of time after blood collection on bacterial survival in human blood

Whole blood was collected and an aliquot (0.9 mL) was immediately taken and mixed, on a 2 mL tube, with 10^5 CFU/mL of bacteria and incubated at 37°C and 80 rpm, for 4h, in a total volume of 1 mL. The remaining blood was kept under the same temperature and agitation conditions for 4h. After this time, a new aliquot of blood was taken and inoculated with 10^5 CFU/mL of bacteria and incubated as before. Bacteria were quantified by CFU counting. This experiment was performed three independent times, using blood from three different donors.

3.2.7 *S. epidermidis* gene expression assays

Three unrelated genes were selected as probes for assessing gene expression stability: *SERP_RS11970*, *SERP_RS10985* and *SERP_RS08870*. Two different experimental set-ups were tested: the utilization of different anticoagulants on the collection tubes and the reduction of the volume of blood used during the co-incubation assays (total incubation volume of 1 mL, 0.6 mL, 0.5 mL, and 0.2 mL). Blood samples were transferred into 2 mL tubes and, then, bacteria were added to each tube to obtain a final concentration of 10^8 CFU/mL. The tubes were incubated for 2h at 37°C and 80 rpm (PSU-10i). After the co-incubation period, samples were sonicated for 5s at 33% amplitude (Cole-Parmer 750- Watt Ultrasonic Homogenizer 230 VAC, IL, USA) to lyse eukaryotic cells. Total RNA isolation, complementary DNA synthesis (cDNA) and quantitative PCR (qPCR) were performed as previously optimized [14], with minor modifications. In brief, after mechanical and chemical lysis of bacterial cells, total RNA was purified using EZNA total RNA kit (Omega Biotek, GA, USA). Genomic DNA was degraded by DNase I (Thermo Scientific, MA, USA) and cDNA synthesized, from 200 ng of total RNA, by RevertAid M-MuLV reverse transcriptase (Thermo Fisher Scientific) and using random primers (NZYTech, Lisboa, Portugal) as priming

strategy. Finally, qPCR was prepared by mixing 2 μL of 1:100 diluted cDNA with 5 μL of Xpert Fast SYBR (Grisp, Porto, Portugal), 0.5 μL of each forward and reverse primers at 0.5 μM and 2 μL of nuclease-free water. The run was completed in a CFX96™ thermal cycler (Bio-Rad, CA, USA) with the following cycling parameters: 3 min at 95°C followed by 40 cycles of 5s at 95°C and 25s at 60°C. The primers used were designed using Primer3 software [15, 16] and synthesized at Metabion (Steinkirchen, Germany). Primers sequences, size of the amplicon and reaction efficiency are presented in Table 3.1. The quantification of the transcripts for each gene under study was determined using 16S rRNA as reference gene and by applying the delta C_q method ($E^{\Delta C_q}$), a variation of the Livak method [17], where $\Delta C_q = C_q$ (reference gene) - C_q (target gene) and E is the experimentally determined reaction efficiency. Reactions efficiency were determined using the dilution method (Pfaffl, 2004) at 60°C.

Table 3.1. List of primers used for the quantification of gene expression by qPCR

Gene	Primer sequence (5′-3')	Product size (base pair)	Efficiency (%)
SERP_RS00125 (16S rRNA)	Fw: GGGCTACACACGTGCTACAA Rv: GTACAAGACCCGGGAACGTA	176	97
SERP_RS11970	Fw: CAGGCATTGAACTTCCCAAT Rv: AATTCGGGGGCATATTTAGG	109	103
SERP_RS10985	Fw: ATGATTTTAGTGCTATCCCTGACT Rv: CACTAATTGCAAGATCATTTTCTG	102	110
SERP_RS08870 (sepA)	Fw: TCTTAAGGCATCTCCGCCTA Rv: GTCTGGTGCGAATGATGTTG	196	97

3.2.8 Statistical analysis

Statistical analysis was carried out with GraphPad Prism Version 6 Trial (CA, USA). For comparisons among different groups one-way or two-away ANOVA, with Tukey´s comparisons test, were used when appropriate (the tests used are detailed in the figure caption). $P < 0.05$ was considered significant.

3.3 RESULTS AND DISCUSSION

Human blood *ex vivo* models have been developed to mimic bloodstream infections as an affordable alternative to *in vivo* models [4, 5, 7]. During human blood collection, anticoagulants need to be used to prevent blood clotting and the choice of the anticoagulant for the *ex vivo* assays may influence the experimental outcome [18, 19]. The most commonly used anticoagulants are heparin, citrate and EDTA [20]. Citrate and EDTA prevent blood from clotting through binding free calcium ions [19], while heparin inhibits coagulation by enhancing the activity of antithrombin III [20]. Initially, pilot experiments were performed to determine the best anticoagulant for *S. epidermidis* gene expression analysis. Thus, the influence of heparin, citrate, and EDTA was evaluated on bacterial growth and transcription levels of the selected genes. As shown in Figure 3.1A, EDTA inhibited bacterial growth in the first 8h, eventually leading to some bacterial death after 24 h of incubation (Figure 3.1A). Furthermore, citrate interferes with calcium levels that may also influence the bacterial growth, depending on the strain, as previously reported [19]. Regarding the gene expression, no significant differences in gene expression were found among the different anticoagulants tested (Figure 3.1B). Taken into consideration these pilot results, the availability and price difference between the anticoagulants tested, and previous experimental results [3, 21] we selected heparin for the rest of the experiments.
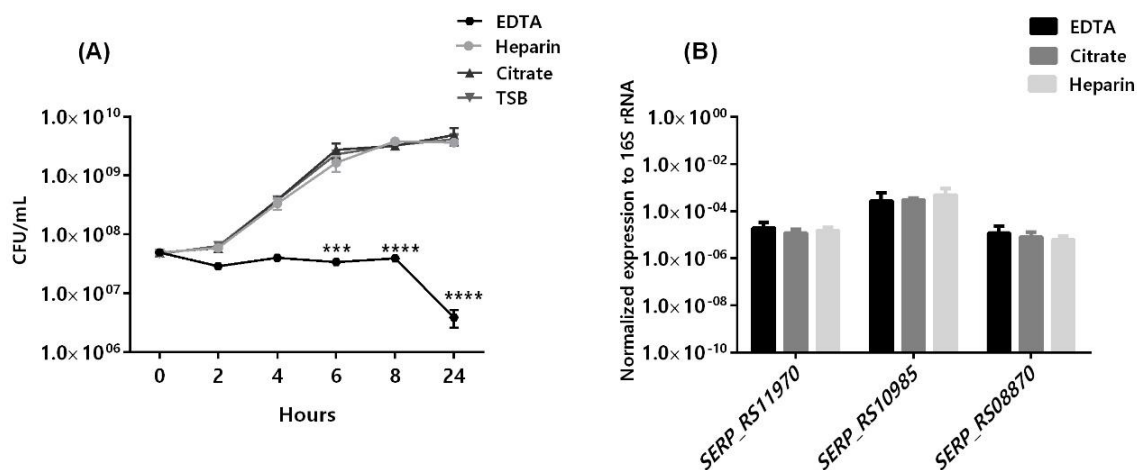
Figure 3.1. The effects of the anticoagulants EDTA, heparin and citrate on *S. epidermidis* PT12003 growth (A) and on transcription levels of *SERP_RS11970*, *SERP_RS10985* and *SERP_RS08870* genes (B). Growth curves are shown in CFU/mL obtained at different time points. The growth curves correspond to the mean ± standard deviation of two independent experiments. Statistical analysis was performed using two-way ANOVA and Tukey´s multiple comparisons test. $***p < 0.001$, $****p < 0.0001$ comparatively to the bacterial growth in TSB. The bars represent the mean plus standard deviation of two independent experiments. Statistical analysis was performed using two-way ANOVA and Turkey´s multiple comparisons test

Another issue related to human blood *ex vivo* models is the bacterial concentration used. Although the quantity of microbes present in human blood during bacteremia is estimated to be up to 10^4 CFU/ mL [22], higher concentrations of bacteria have been used in human blood *ex vivo* models [4, 23, 24], due to the lack of sensitivity of many experimental methods to assess lower stimuli [25, 26]. For instance, it is known that for the analysis of the transcriptomic response, the initial bacterial concentration needs to be significantly higher, to ensure a sufficient amount of RNA for downstream applications [27]. Since this is one of our goals in this thesis, we then investigated the effect of different initial bacterial concentration (10^4-10^8 CFU/mL) on the ability of *S. epidermidis* to survive in human blood.
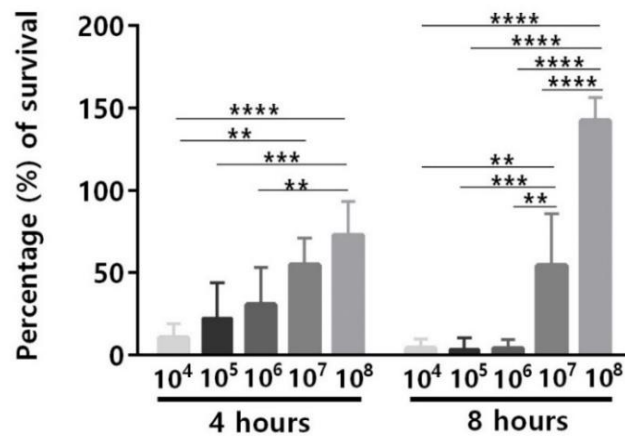
Figure 3.2. The effect of initial bacterial concentrations on the ability of *S. epidermidis* to survive in human blood after 4h and 8h of incubation. The bars represent the mean plus standard deviation of five independent experiments, performed with five different donors. Statistical analysis was performed using two-way ANOVA and Tukey´s multiple comparisons test. Significant differences are depicted with **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Figure 3.2 represents the percentage of bacterial survival after 4 and 8h of co-incubation. Since it is known that there is a significant source of experimental variability when working with human samples, due to the is inherent to particular traits of the donors such as age [28], gender [29] and the proportion of different blood cell populations [30], this experiment was performed with blood from five different donors to increase the significance of our results. Not surprisingly, the ability of *S. epidermidis* to survive in human blood was cell concentration dependent: the lower the inoculum, the higher the percentage of bacterial killing by human blood (Figure 3.2). Interestingly, after 4h of co-incubation, cell death was observed in all tested bacterial concentrations but after

8h of co-incubation, the higher concentration inocula presented higher cell density than at time zero, suggesting that under these conditions, blood had no antimicrobial activity and bacteria were able to grow.

Based on this latter result, it was hypothesized the increase of bacteria survival, in the higher inoculum experiment, was related to the loss of viability of blood cells. Therefore, the viability of leukocytes was analyzed by flow cytometry, immediately, 4 and 8h after blood collection as is shown in Figure 3.3.
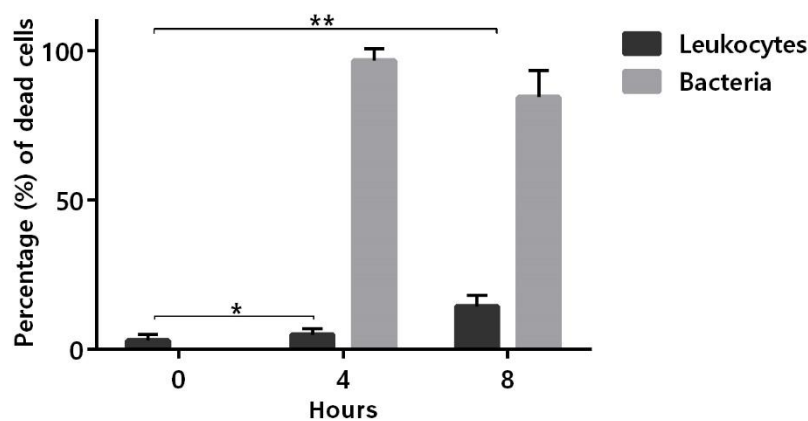
Figure 3.3. The percentage of dead leukocytes and bacterial cells after 4 and 8h of blood collection. Leukocytes viability was assessed by flow cytometry immediately, 4 and 8h after blood collection. Bacteria were added to blood immediately and 4h after its collection. In all time points, the co-incubations assays were performed for 4h. The bars represent the mean plus standard deviation of three independent experiments, performed using three different donors. Statistical analysis was performed using one-way ANOVA and Turkey´s multiple comparison test. $*p < 0.05$. $**p < 0.01$.

The results demonstrated that right after blood collection, 3% of leukocytes were already dead. This fact may be related to the process of collection and processing time of human blood [31]. The results also showed that 4h after blood collection, this percentage slightly raised to 5% and then to 15% after 8h. Since higher antimicrobial activity of human blood was observed during the first 4 h of incubation, and the loss of leukocytes viability was significantly lower, these data suggested that prolonged incubation periods can negatively influence the experimental outcome, particularly if a higher bacterial inoculum is needed, such as when performing RNA-sequencing [4, 5, 32, 33]. Nolte and colleagues developed a modified *ex vivo* model to study later aspects of septicemia caused by *Neisseria meningitidis* using 12 and 24h of the co-incubation period in human blood [6]. However, as shown by these results, this an extensive period of incubation could yield a higher number of dead leukocytes and, as such, should be avoided when determining bacterial survival

studies in human blood. The loss of leukocytes viability can occur when research laboratories are not physically close to human blood collecting centers, and this should also be taken into consideration when planning experiments using human blood. An easy way to overcome this obstacle is to maintain the blood under gentle agitation [34].

To determine if the observed loss of leukocytes viability has an impact on the ability of bacteria to survive, co-incubation assays were also performed but controlling the time between blood collection and blood utilization. For this assay, we selected a bacterial inoculum of 10^5 CFU/mL, based on our previous results (Figure 3.2), and taken into consideration that this concentration is more closely related to the maximum *in vivo* bacterial load [22]. As can be seen in Figure 3.3, the results showed that a higher antimicrobial activity was detected when blood was used immediately after collection, resulting in 97% of bacterial cell death. In contrast, when blood was used after 4 h of its collection, only 84% of bacterial death was detected. Interestingly, a 2% reduction in leukocyte viability (from 3 to 5%) resulted in 13% reduction in antimicrobial activity (from 97% to 84%). This confirms that blood seems to lose its antimicrobial properties when very long periods of time occur between blood collection and utilization, in this case, 4h of interval. This fact is important to consider when planning *ex vivo* experiments with blood, as it may have substantial consequences on the results obtained.

Another important practical aspect when considering using human blood as an *ex-vivo* model is the limitation of blood availability. As such, the ability to reduce the volume of blood per experiment, without compromising the results, is of interest. When analyzing several gene expression studies published, it was observed that different volumes of human blood, ranging from 0.2 mL to 80 mL per experiment, have been used [4, 7, 8, 12, 27]. As the long-term goal of the research group is to study *S. epidermidis* gene expression in human blood, there was interest in determining if reducing the volume of blood in the co-incubation assays down to 0.2 mL had a detrimental effect on the stability of gene transcription. The starting volume was 0.9 mL, in order to compare with our previous data obtained using RNA-seq [23]. To do so, the transcription of 3 unrelated genes was assessed. One of the selected genes was *sepA* (*SERP_RS08870*), which codifies a protease that plays an important role in bacterial immune invasion through the degradation of AMPS produced by the host [35]. The two other selected genes were *SERP_RS11970*, a gene that codifies a major facilitator superfamily and *SERP_RS10985*, a universal stress protein.

As shown in Figure 3.4A, no significant differences were found in the expression of the selected genes using, in any of the different volumes of human blood tested. Noteworthy, a volume of blood as low as 0.18 ml per reaction can be used without impacting the transcription of the selected genes. Interestingly, we also observed that the reduction of blood volume did not affect the percentage of bacterial survival, as shown in figure 3.4B.
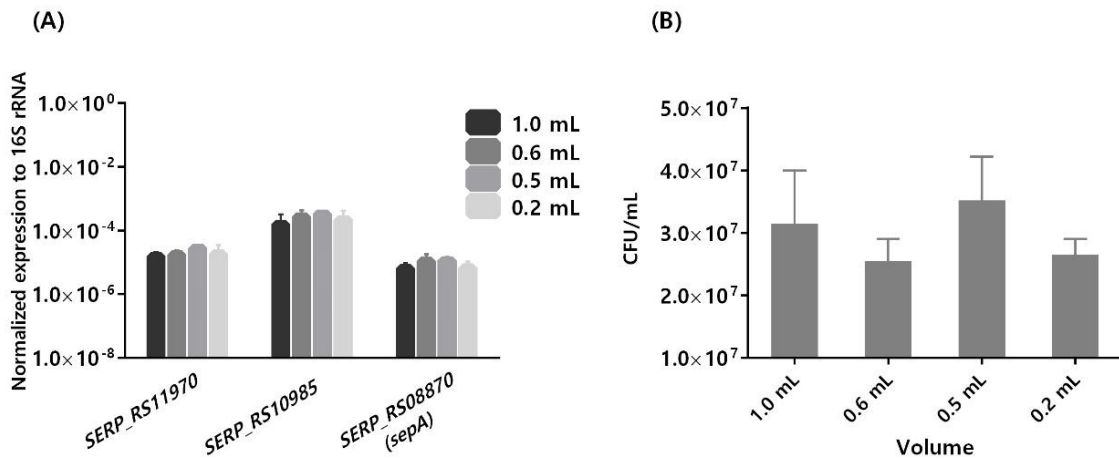
Figure 3.4. The influence of using different volumes of human blood in co-incubations assays on the stability of transcription levels of *SERP_RS11970*, *SERP_RS10985* and *SERP_RS08870* genes (A) and on bacterial cells culturability (B). The bars represent the mean plus standard deviation of three independent experiments, performed with three different donors. Statistical analysis was performed, respectively, using two-way or one-way ANOVA and Tukey´s multiple comparisons test

3.4 CONCLUSION

There remains a great deal of work to be done in clarifying the factors that contribute to *S. epidermidis* survival and adaptation to blood, which contribute to the evasion from the host immune system. The improvement of the human blood *ex vivo* model will help to plan a better and more reliable model and, consequently, to obtain results with higher clinical relevance. These findings are of technical importance for future studies since it highlights key parameters that should be considered when using human blood as an *ex vivo* model for the analysis of the gene expression of *S. epidermidis,* in particular, the possibility of using low volume of blood per reaction, without compromising the experimental results, at least in regard to the parameters tested herein.

3.5 REFERENCES

1. **Otto M**. *Staphylococcus epidermidis*–the 'accidental' pathogen. *Nature Reviews Microbiology* 2009;7(8):555-567.

2. **Loza-Correa. M, Juan A. Ayala, Iris Perelman, Keith Hubbard, Miloslav Kalab** *et al.* The peptidoglycan and biofilm matrix of *Staphylococcus epidermidis* undergo structural changes when exposed to human platelets. *PLoS ONE* 2019;14(1):e0211132.

3. **França A, Carvalhais V, Maira-Litrán T, Vilanova M, Cerca N** *et al.* Alterations in the *Staphylococcus epidermidis* biofilm transcriptome following interaction with whole human blood. *Pathogens and Disease* 2014;70(3):444-448.

4. **Qin L, Da F, Fisher EL, Tan DC, Nguyen TH** *et al.* Toxin Mediates Sepsis Caused by Methicillin-Resistant *Staphylococcus epidermidis. PLoS Pathogens* 2017;13(2):e1006153.

5. **Malachowa N, Whitney AR, Kobayashi SD, Sturdevant DE, Kennedy AD** *et al.* Global Changes in *Staphylococcus aureus* Gene Expression in Human Blood. *PLoS ONE* 2011;6:e18617.

6. **Nolte O, Rickert A, Ehrhard I, Ledig S, Sonntag HG**. A modified ex vivo human whole blood model of infection for studying the pathogenesis of *Neisseria meningitidis* during septicemia. *FEMS Immunology Medical Microbiology* 2002;32(2):91-95.

7. **Echenique-Rivera H, Muzzi A, Del Tordello E, Seib KL, Francois P** *et al.* Transcriptome analysis of *Neisseria meningitidis* in human whole blood and mutagenesis studies identify virulence factors involved in blood survival. *PLoS Pathog* 2011;7(5):e1002027.

8. **Mereghetti L, Sitkiewicz I, Green NM, Musser JM**. Extensive Adaptive Changes Occur in the Transcriptome of *Streptococcus agalactiae* (Group B Streptococcus) in Response to Incubation with Human Blood. *PLoS ONE* 2008;3:e3143.

9. **Vebo HC, Snipen L, Nes IF, Brede DA**. A modifed ex vivo human whole blood model of infection for studying the pathogenesis of *Neisseria meningitidis* during septicemia. *PLOS ONE* 2009;4(11).

10. **Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C** *et al.* Stage-specific gene expression of *Candida albicans* in human blood. *Mol Microbiol* 2003;47(6):1523-1543.

11. **Fradin C, De Groot P, MacCallum D, Schaller M, Klis F** *et al.* Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood: Neutrophils trigger *C. albicans* response. *Molecular Microbiology* 2005;56:397-415.

12. Hunniger K, Lehnert T, Bieber K, Martin R, Figge MT *et al.* A virtual infection model quantifies innate effector mechanisms and *Candida albicans* immune escape in human blood. *PLoS computational biology* 2014;10(2):e1003479.

13. Freitas AI, Lopes N, Oliveira F, Bras S, Franca A *et al.* Comparative analysis between biofilm formation and gene expression in *Staphylococcus epidermidis* isolates. *Future Microbiology* 2018;13:415-427.

14. França A, Freitas AI, Henriques AF, Cerca N. Optimizing a qPCR Gene Expression Quantification Assay for *S. epidermidis* Biofilms: A Comparison between Commercial Kits and a Customized Protocol. *PLoS ONE* 2012;7:e37480.

15. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC *et al.* Primer3–new capabilities and interfaces. *Nucleic Acids Research* 2012;40(15):22.

16. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 2007;23(10):1289-1291.

17. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2-\Delta\Delta CT$ Method. *Methods* 2001;25:402-408.

18. Freitas M, Porto G, Lima JL, Fernandes E. Isolation and activation of human neutrophils in vitro. The importance of the anticoagulant used during blood collection. *Clinical Biochemistry* 2008;41(7-8):570-575.

19. Strobel L, Johswich KO. Anticoagulants impact on innate immune responses and bacterial survival in whole blood models of Neisseria meningitidis infection. *Scientific reports* 2018;8(1):10225.

20. Engstad CS, Gutteberg TJ, Osterud B. Modulation of blood cell activation by four commonly used anticoagulants. *Thrombosis and haemostasis* 1997;77(4):690-696.

21. Franca A, Pier GB, Vilanova M, Cerca N. Transcriptomic Analysis of *Staphylococcus epidermidis* Biofilm-Released Cells upon Interaction with Human Blood Circulating Immune Cells and Soluble Factors. *Frontiers Microbiology* 2016;7(1143).

22. Opota O, Croxatto A, Prod'hom G, Greub G. Blood culture-based diagnosis of bacteraemia: state of the art. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2015;21(4):313-322.

23. França A, Carvalhais V, Maira-Litrán T, Vilanova M, Cerca N *et al.* Alterations in the *Staphylococcus epidermidis* biofilm transcriptome following interaction with whole human blood. *Pathogens and Disease*.

24. Askarian F, Uchiyama S, Valderrama JA, Ajayi C, Sollid JU *et al.* Serine-Aspartate Repeat Protein D Increases *Staphylococcus aureus* Virulence and Survival in Blood. *Infection and Immunity* 2017;85(1).

25. Machado A, Almeida C, Carvalho A, Boyen F, Haesebrouck F *et al.* Fluorescence in situ hybridization method using a peptide nucleic acid probe for identification of *Lactobacillus spp.* in milk samples. *International journal of food microbiology* 2013;162(1):64-70.

26. Bacconi; A, Gregory S. Richmond, Michelle A. Baroldi, Thomas G. Laffler, Lawrence B. Blyn *et al.* Improved Sensitivity for Molecular Detection of Bacterial and *Candida* Infections in Blood. *Journal of Clinical Microbiology* 2014;52(9):3164–3174.

27. Hedman AK, Li MS, Langford PR, Kroll JS. Transcriptional profiling of serogroup B *Neisseria meningitidis* growing in human blood: an approach to vaccine antigen discovery. *PLOS ONE* 2012;7(6):e39718.

28. Eady JJ, Wortley GM, Wormstone YM, Hughes JC, Astley SB *et al.* Variation in gene expression profiles of peripheral blood mononuclear cells from healthy volunteers. *Physiol Genomics* 2005;22(3):402-411.

29. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC *et al.* Individuality and variation in gene expression patterns in human blood. *Proceedings of the National Academy of Sciences* 2003;100(4):1896-1901.

30. Cobb JP, Mindrinos MN, Miller-Graziano C, Calvano SE, Baker HV *et al.* Application of genome-wide expression analysis to human health and disease. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(13):4801-4806.

31. Ferrante A, Thong YH. Optimal conditions for simultaneous purification of mononuclear and polymorphonuclear leucocytes from human blood by the Hypaque-Ficoll method. *Journal Immunology Methods* 1980;36(2):109-117.

32. França A, Pier GB, Vilanova M, Cerca N. Transcriptomic Analysis of *Staphylococcus epidermidis* Biofilm-Released Cells upon Interaction with Human Blood Circulating Immune Cells and Soluble Factors. *Frontiers in Microbiology*.

33. Franca A, Cerca N. Plasma is the main regulator of *Staphylococcus epidermidis* biofilms virulence genes transcription in human blood. *Pathogens and Disease* 2016;74(2):27.

34. Afonso G, Scotto M, Renand A, Arvastsson J, Vassilieff D *et al.* Critical parameters in blood processing for T-cell assays: validation on ELISpot and tetramer platforms. *Journal Immunology Methods* 2010;359(1-2):28-36.

35. **Lai Y, Villaruz AE, Li M, Cha DJ, Sturdevant DE** *et al.* The human anionic antimicrobial peptide dermcidin induces proteolytic defence mechanisms in staphylococci. *Molecular Microbiololgy* 2007;63(2):497-506.

CHAPTER 4

In silico analysis of the transcriptome of clinical and commensal *S. epidermidis* isolates upon interaction with human blood

SUMMARY

The current inability to discriminate between true bacteremia caused by *S. epidermidis* and contaminated blood cultures can delay or mislead diagnosis, increasing morbidity and treatment costs. Hence, the main goal of the study described in this chapter was to identify possible RNA-based molecular markers for the diagnosis of *S. epidermidis* infections. For that, the transcriptome of three clinical and three commensal isolates from the community was sequenced using RNA-sequencing technology (RNA-seq). The comparison of the six transcriptomes revealed that the gene *SERP_RS11970* could be a potential molecular marker able to discriminate clinical from commensal isolates.

4.1 BRIEF INTRODUCTION

S. epidermidis is a normal inhabitant of human skin but has become an important nosocomial pathogen that causes bacteremia in patients with indwelling devices, as previously described in Chapter 2. Blood cultures carried out for diagnosis purposes can be easily contaminated by *S. epidermidis* living on the skin and mucosae. The current inability to discriminate between true bacteremia caused by *S. epidermidis* and contaminated blood cultures often leads to misdiagnosis, resulting in a significant increase in patient morbidity and in health care costs [1]. A fast and accurate diagnosis is, therefore, essential to reduce the clinical and economic impact associated with *S. epidermidis* infections.

Several DNA-based molecular strategies attempting to unravel possible differences in genetic profiles of invasive and contaminant strains have been tested before, including both genome-wide [2, 3] and gene specific approaches [4, 5] but, so far, no RNA-based strategy has been attempted. Since gene expression changes in response to environmental conditions, differential expression of specific genes can determine the ability of *S. epidermidis* to adapt to an environment like human blood and, eventually, to cause infection. Therefore, a transcriptomic approach may be more indicated to identify differences between commensal and clinical isolates [6].

RNA-seq has become the most common technique to analyze bacterial transcriptomes [7]. It is a high-throughput sequencing technology that allows us the entire transcriptome to be analyzed in a quantitative manner [8]. This methodology enables to measure the mRNA expression levels that can be quantified as RPKM (reads per kilobase per million mapped reads) values [9]. RNA-seq has several advantages in comparison with other transcriptomic technologies such as microarrays, including higher sensitivity [10], higher technical reproducibility [11] and a larger dynamic range for measuring variability in expression levels [9], consequently, it is much more discriminatory [12]. Moreover, a reference sequenced genome is not required [7].

The aim of the study described in this chapter was to analyze the transcriptome of three *S. epidermidis* clinical isolates and three commensal isolates, previously obtained by RNA-seq, in order to identify possible molecular markers for the diagnosis of *S. epidermidis* infections.

4.2 MATERIAL AND METHODS

4.2.1 Biological model used to obtain the transcriptomic data for *in silico* analysis

The raw data of the transcriptome of *S. epidermidis* isolates were previously obtained in the scope of the project FCOMP-01-012014-FEDER-041246 (EXPL/BIAMIC/0101/2013). Three clinical and three commensal isolates were tested (Table 4.1). A human blood *ex vivo* model was used, as previously described [13]. Briefly, biofilm released cells of each isolate, obtained as described before [14], were adjusted to the same concentration (10^8 CFU/mL) and co-incubated at 37°C for 2h in human blood, collected from adult healthy volunteers. RNA was then extracted, and the quality analyzed. Only samples with RNA quality numbers superior to 9 were used to construct the cDNA libraries. Sequencing was performed using a MiSeq® sequencer (Illumina) from paired-end reads (2 × 150 bp).

Table 4.1. List of *S. epidermidis* isolates used in this study

Strain	Description	Reference
IE214	Clinical isolate from a patient with infective endocarditis	[15]
PT12003	Clinical isolate from a patient after stomach surgery	[16]
1457	Central venous catheter infection	[17]
SECOM005A	Commensal isolate from a healthy individual	
SECOM020A1	Commensal isolate from a healthy individual	[18]
SECOM030A	Commensal isolate from a healthy individual	

4.2.2 RNA-sequencing data analysis

The alignment of the nucleic acid sequencing reads, using *S. epidermidis* RP62A (GenBank accession number: CP000029.1) as the template, the normalization of the reads *per* kilobase *per* million mapped reads (RPKM) [9] and the analysis of differential gene expression were carried out using CLC Genomics Workbench, version 5.5.1.

In order to try to discriminate clinical from commensal isolates, an *in silico* comparative transcriptomic analysis was performed. By using Venn diagrams (version 2.1) [19], genes that were detected in clinical or commensal isolates at determined RPKM thresholds and genes detected in both groups of isolates were identified. This analysis was based on RPKM threshold and three strategies were devised, in order to maximize the potential number of genes of interest (Figure 4.1). The first strategy was performed to identify genes only expressed in clinical or in

commensal isolates. RNA-sequence raw data with RPKM <1 was trimmed and discarded, taken into consideration the sensitivity and detection limit of the method [9, 20]. Thus, using a Venn diagram, genes that were commonly expressed in all clinical isolates were selected. These genes were then contra-selected against the three commensal isolates gene expression profiles. Therefore, genes expressed only in the clinical isolates group were selected for further analysis. A similar analysis was performed by first selecting the commonly expressed genes in all three commensal isolates and then contra-selecting against each of the clinical isolates transcriptomes.

Analysis RNA-seq data

1	2	3
RPKM >1	RPKM >10 000	RPKM >1 (all groups) RPKM >10 RPKM >100 RPKM >1000 RPKM >5000
Select genes expressed in a specific group	Select genes expressed at different levels in all groups	Select genes expressed at different levels in a specific groups and all groups
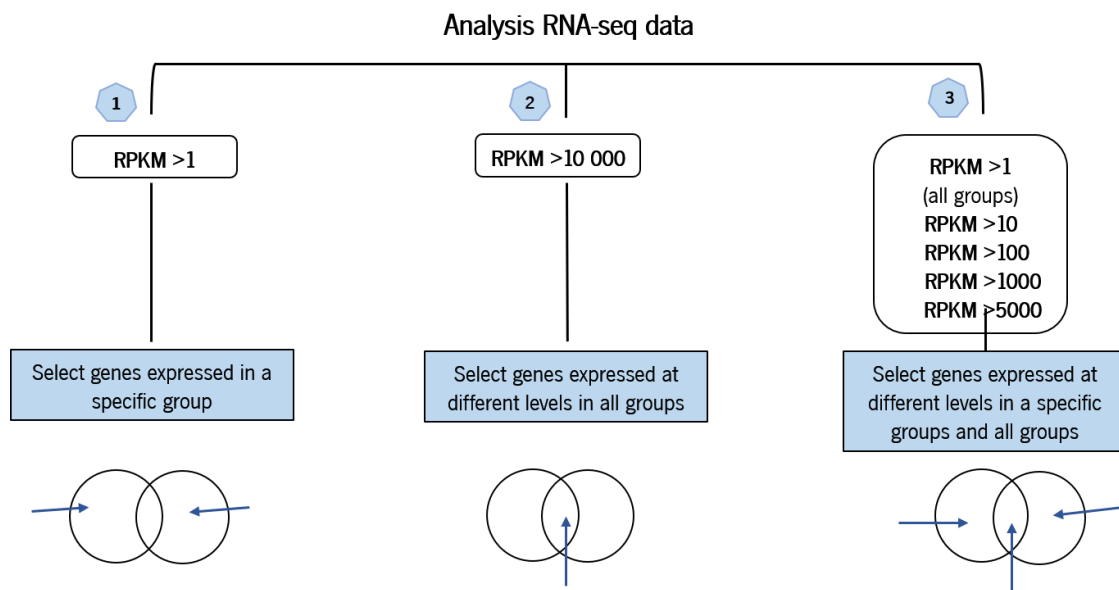
Figure 4.1. The workflow of RNA-seq data analysis. The bioinformatics analysis based on RPKM thresholds was performed using three strategies. The first strategy identified genes only detected in either clinical or commensal isolates. The second and third strategies were performed to identify differentially expressed genes using different RPKM thresholds (RPKM >1, RPKM >10, RPKM >100, RPKM >1000, and RPKM >5000).

A second strategy was performed aiming to identify commonly highly expressed genes (RPKM >10 000) in all isolates, but with different expression levels between the two groups, in order to enable the discrimination between the two groups. Similarly, to what was done in the first analysis, Venn diagrams were generated to select genes that were expressed in all clinical isolates with a minimum RPKM value of 10 000. These genes were then contra-selected against the three commensal isolates gene expression profiles. The reverse analysis was also performed.

Due to the low number of potential genes highlighted in the first two strategies, a third approach was pursued, where the differential expressed genes found in any of the groups were clustered using increasing threshold values, including RPKM >10, RPKM >100, RPKM >1000 and RPKM

>5000. Venn diagrams were generated using the different RPKM thresholds and preliminary lists of potential genes of interest were created by determining the genes that were present in the lower RPKM threshold but absent in the next tier of RPKM threshold. The theoretical discriminative power of each gene was determined by dividing the average of RPKM values of the clinical group by the average of the commensal group. Genes of potential interest were considered when (i) the theoretical fold difference was above 2 or (ii) when it was above 4, for genes that were expressed at a low level (since there are a higher number of genes at lower RPKM values).

For all potential genes of interest, the Search Tool for the Retrieval of Interacting Genes/proteins (String) version 9.05 [21] was used in order to determine the predicted function of the selected genes. In addition, UniProt repository [22] was also used to determine the predicted function of proteins that were not identified by STRING.

4.2.3 RNA-seq validation

To confirm the results obtained by *in silico* analysis, both technical and biological validations were performed. Technical validation was performed using the same total RNA utilized in the cDNA library construction. The biologic validation was performed using the same isolates and the same experimental conditions used in RNA-seq, but with different human blood donors, as detailed bellow.

4.2.3.1 Bacterial Strains and Growth conditions

Growth conditions used herein were the same used for RNA-seq analysis, following a protocol previously optimized [14]. Briefly, one *S. epidermidis* colony was inoculated into 2 mL of Tryptic Soy Broth (TSB) (Liofilchem, Teramo, Italy) and incubated overnight at 37∘C and with agitation at 120 rpm (ES-20 Shaker-Incubator, Biosan, Riga, Latvia). The overnight culture was adjusted to an optical density (OD), at 640 nm, of 0.250 ± 0.05 with TSB. Biofilms were formed by inoculating 15 μL of this suspension into 1 mL of TSB supplemented with 0.4% (v/v) glucose (TSB$_G$) and incubated in a 24-well plate (Orange Scientific, Braine-l'Alleud, Belgium) at 37∘C and 120 rpm for 24h. After this period, the spent medium was carefully removed and biofilms were washed twice with 0.9 % NaCl, followed by careful addition of 1 mL of fresh TSB$_G$ and subsequent incubation at the same temperature and agitation conditions. Biofilm-released cells were collected from 12 different originating biofilms and pooled together to decrease the variability inherent to biofilm growth [23]. The suspensions were then sonicated for 10s, at 33% amplitude (Cole- Parmer 750 –Watt Ultrasonic Homogenizer 230 VAC, IL, USA), in order to dissociate cell clusters. This

sonication cycle did not affect cell viability [24]. Thereafter, the concentration of cells was adjusted to 1×10^9 total CFU/mL, before using in the blood co-incubation assays.

4.2.3.2 Human blood collection

Peripheral blood was collected from healthy adult volunteers, not taking antibiotics or anti-inflammatory medication within the last 14 days, by venipuncture into BD Vacutainer® tubes spray coated with lithium heparin (Becton Dickinson, NJ, USA). Blood was collected under a protocol approved by the Institutional Review Board of the University of Minho (SECVS 002/2014 (ADENDA)), which is in strict accordance with the Declaration of Helsinki and Oviedo Convention. All donors gave written informed consent to have blood taken.

4.2.3.3 Co-incubation of bacteria with human blood

In 2 mL tubes, 100 μL of a suspension of 1×10^9 CFU/mL was mixed with 900 μL of human blood and incubated at 80 rpm (PSU-10i, Biosan), for 2h at 37 °C. After the co-incubation period, in order to lyse eukaryotic cells and, this way, get rid of eukaryotic nucleic acids, the samples were sonicated for 5s at 33% amplitude. Then, bacteria were harvested by 5 min centrifugation at 16000 g at 4 °C. This assay was performed three independent times using blood of three different donors.

4.2.3.4 RNA extraction

RNA was isolated as previously optimized [25], combining mechanical lysis with silica membrane-based RNA isolation, using EZNA total RNA kit (Omega Biotek, Norcross, GA). Briefly, the bacterial pellet was suspended in 500 μL of TRK lysis buffer (supplemented with 20 μL/mL of β-mercaptoethanol plus 500 μL of phenol). The resulting suspension was transferred into 2 mL safe lock tubes containing 0.5 g of acid-washed 150–21 2 mm silica beads (Sigma, MO, USA). The tubes were then placed into a cell disruptor (FastPrep®-24, MP Biomedicals, CA, USA) and run for 35s at 6.5 m/s. The samples were immediately placed on ice for 5 min and the beat-beading step was repeated thrice. Afterwards, samples were centrifuged at 16000 g for 1.5 min, the supernatants transferred into 2 mL DNase/ RNase-free tubes and mixed with an equal volume of 70% ethanol. The samples (including any remaining precipitate) were transferred into the silica-membrane columns and centrifuged at 15000 g for 1 min at room temperature (RT). The following steps were performed according to the manufacturer's instructions.

4.3.2.5 DNase treatment

To degrade contaminating genomic DNA, each RNA sample was digested with 2 µL of DNase I plus 4 µL of 10 × DNase I Reaction Buffer (Thermo Scientific, MA, USA), and incubated at 37 °C for 30 min. Then, to inactivate the DNase I activity, 4 µL of 25 mM EDTA (pH 8.0) were added to the mixture and incubated, at 65 °C, for 10 min.

4.3.2.6 RNA quality determination

RNA integrity was assessed by visualization of the 23S/16S rRNA band pattern. RNA samples were analyzed in a 1% (w/v) agarose gel. Non-denaturing electrophoresis was carried out at 80 V for 60 min. The gel was stained with Midori Green DNA staining (Nippon Genetics Europe GmbH, Germany) in Tris–acetate–EDTA (TAE) buffer and visualized using ChemiDoc™ XRS+ (Bio-Rad, CA, USA). RNA concentration and purity were determined using a NanoDrop™ 1000 spectrophotometer (Thermo Scientific, MA, USA).

4.3.2.7 cDNA synthesis

cDNA was synthesized, from 200 ng of total RNA, by RevertAid M-MuLV reverse transcriptase (Thermo Scientific, MA, USA) and using random primers (NZYTech, Lisboa, Portugal), as priming strategy. To determine the possibility of genomic DNA carry-over, control reactions were performed under the same conditions but lacking the reverse transcriptase enzyme (NRT control).

4.3.2.8 Quantitative PCR (qPCR)

In order to validate RNA-seq data, qPCR was performed to quantify the transcription of selected genes using (i) the same total RNA utilized for libraries construction (technical validation) and (ii) total RNA from independent experiments performed under the same biological conditions (biological validation). Primers for each selected gene were designed using Primer3 software [26] having *S. epidermidis* RP62A as the template (Table 4.2). Finally, qPCR was prepared by mixing 2 µL of cDNA, diluted 1:100, with 5 µL of Xpert Fast SYBR (Grisp, Porto, Portugal), 0.5 µL of each forward and reverse primers at 0.5 µM and 2 µL of nuclease-free water. qPCR was performed using a CFX96™ with the following cycling parameters: 3 min at 95 °C followed by 39 cycles of 5s at 95 °C and 25s at 60 °C. The quantification of the specific transcripts for each gene under study was determined using 16S rRNA as reference gene and by applying the delta C_q method ($E^{\Delta C_q}$), a variation of the Livak [27] method, where $\Delta C_q = C_q$ (reference gene) - C_q (target gene) and E is the experimentally determined reaction efficiency. Reaction efficiency was determined using the dilution method [28] at 60 °C.

4.2.4 Statistical analysis

Statistical analysis was carried out with GraphPad Prism Version 6 Trial (CA, USA). $p < 0.05$ was considered significant.

Table 4.2. List of primers used, for qPCR, to determine the level of transcription of the genes of interest

Gene	Primer sequence (5´-3')	Product size (base pair)	Efficiency (%)
SERP_RS00125 (16S rRNA)	Fw: GGGCTACACACGTGCTACAA Rv: GTACAAGACCCGGGAACGTA	176	97
SERP_RS10985	Fw: ATGATTTTAGTGCTATCCCTGACT Rv: CACTAATTGCAAGATCATTTTCT	102	111
SERP_RS11970	Fw: CAGGCATTGAACTTCCCAAT Rv: AATTCGGGGGCATATTTAGG	109	103
SERP_RS11770	Fw: ATAACCAAATGCGTGAGCAA Rv: ACAGCAGCACAAGCAGAAGA	105	94
SERP_RS00240	Fw: TCACAAGAATTTTTGGCTGAG Rv: TTGATTTGCACGTTTTTCAA	167	98
SERP_RS10265	Fw: ATTCGTGTCAGAAACCGCTCAA Rv: GCATTTCCAGCGTTTCCTTCAT	102	91
SERP_RS11115	Fw: CACAATCAATTCCCGCAAGT Rv: GACGGCATCACCACTTCCTA	182	108
SERP_RS04470	Fw: TTTCCTGATTGGCTTCAGTC Rv: GCCAAACGCTTCTATATCCA	108	107
SERP_RS06815	Fw: CCCACTGGAGCTCAATCAGAATTA Rv: AAACTCTTCGCCAATTGAACTTCC	132	92
SERP_RS05305	Fw: AATGTACAGCCCGGTCCAATAGAT Rv: TATTTGCTTGGCAAGTTGTTGAGG	117	99

Table 4.2. Continued

SERP_RS04475	Fw: AAAGTGCGGCTAAATCATTAAA	133	108
	Rv: TGATGCTCATGTACCTCCAA		

4.3 RESULTS AND DISCUSSION

4.3.1 Transcriptomics analysis

The transcriptome of three clinical and three commensal isolates after interaction with human blood was obtained by RNA-seq. In order to select potential markers that allow the discrimination between clinical and commensal isolates, bioinformatics analysis was performed using different RPKM thresholds. Venn diagrams were used to display the number of common and different genes expressed in clinical or commensal isolates. For comparative analysis, three different strategies were performed. The first approach used was to select genes expressed only in one of the groups (clinical or commensal isolates), trimming data with RPKM <1. Under this analysis, three genes uniquely expressed in the clinical group were found (Figure 4.2A). The same process was also performed but starting the analysis with the common genes expressed in the commensal isolates and then compared to the clinical isolates, where two genes were detected (Figure 4.2B).

Figure 4.2. Venn diagrams showing the number of common transcripts (overlapping circles) and unique transcripts (non-overlapping circles) in clinical (A) and commensal isolates (B) using RPKM >1 as threshold.

As shown in Table 4.3, of the five identified genes, four had very low levels of expression, being *SERP_RS11970* the exception. However, it should be noted that may be due to the artifacts possibly created during the sequencing process, the non-detection of a transcript by RNA-seq does not always mean that the gene is not being expressed [29].

For this reason, RNA-seq data shall always be confirmed. Furthermore, taken into consideration the sensitivity of this method, ideally, genes with bigger differences between RPKM values should be the first choice for downstream analysis.

Table 4.3. RPKM values of the genes selected through the first strategy of analysis, which were uniquely expressed in each specific group

Genes	Clinical isolates			Commensal isolates		
	1457	IE214	PT12003	SECOM005A	SECOMO020A1	SECOM030A
SERP_RS04415	28.55	5.92	13.13	<1	<1	<1
SERP_RS12610	18.58	17.98	19.93	<1	<1	<1
SERP_RS03045	5.26	6.96	9.5	<1	<1	<1
SERP_RS12010	<1	<1	<1	1.78	3.51	3.52
SERP_RS11970	<1	<1	<1	148.8	170.3	81.75

Because the number of genes detected in this first analysis was very reduced, with only one top candidate to pursue in the follow-up experiments, a second strategy was performed in order to select highly expressed genes in both groups, but which relative expression could potentially differentiate clinical from commensal isolates. As shown in Figure 4.3A, no highly expressed gene was uniquely detected in any of the groups, but six genes were common to all isolates.

Figure 4.3. Venn diagrams showing the number of common transcripts (overlapping circles) and unique transcripts (non -overlapping circles) in clinical (A) and commensal isolates (B) using RPKM >10 000 as threshold.

Contrary to what was shown in the first analysis, it was indifferent to first select highly expressed genes in the clinical or in the commensal isolates. Unfortunately, as shown in Table 4.4, the RPKM values obtained for these genes were similar in both groups of isolates, suggesting that these are

core genes used by both commensal and clinical isolates when in contact in human blood and, as such, are not useful to discriminate the two groups of isolates.

Table 4.4. RPKM values of the genes identified through the second strategy of analysis, using a RPKM > 10 000 as a threshold

Genes	Clinical isolates			Commensal isolates		
	1457	IE214	PT12003	SECOM005A	SECOM020A1	SECOM030A
SERP_RS00780	61706.11	78045.10	57830.02	64871.65	34332.44	66740.57
SERP_RS02950	18548.44	33706.33	10886.29	32380.24	17601.57	19390.07
SERP_RS04865	64631.33	38311.95	31002.46	49439.44	56082.85	65545.68
SERP_RS12515	51987.01	45481.55	49682.37	36675.02	35324.47	42636.32
SERP_RS12505	676005.7	387005.2	365940.8	261519.34	378181.2	701467.95
SERP_RS12510	68989.32	26861.08	36297.75	31765.74	89267.09	41567.03

Taken into consideration the previous results, a third strategy was performed, aiming to identify other genes differentially expressed between both groups, but with sufficient discriminative power. In order to do so, several Venn diagrams were constructed using different RPKM threshold (RPKM >1 (Figure 4.1), RPKM >10, RPKM >100, RPKM >1000 and RPKM >5000 (Figure 4.4).

As described above, the genes identified when using a threshold of RPKM >1 are listed in Table 4.3. At a RPKM >10, six and four genes only detected in clinical or commensal isolates were found, respectively. On the other hand, at RPKM >100 threshold, seven genes were obtained in clinical isolates and three genes in commensal isolates. Moreover, at RPKM >1000 threshold only one gene was detected in commensal isolates (Table 4.5).
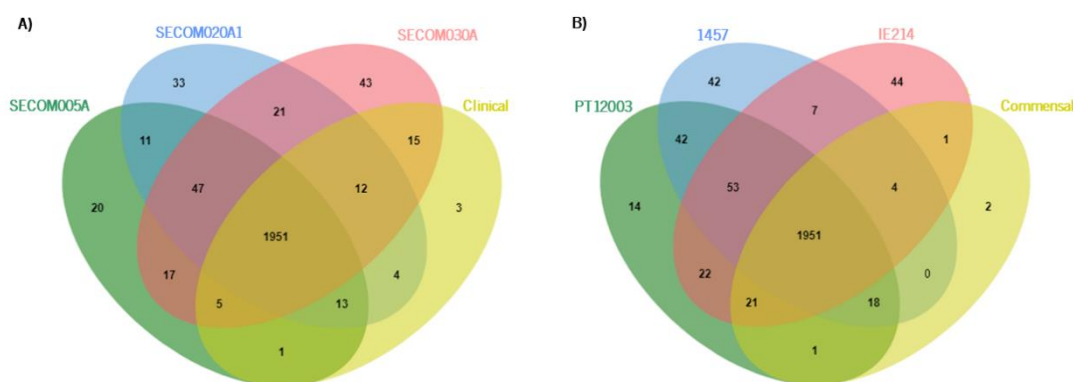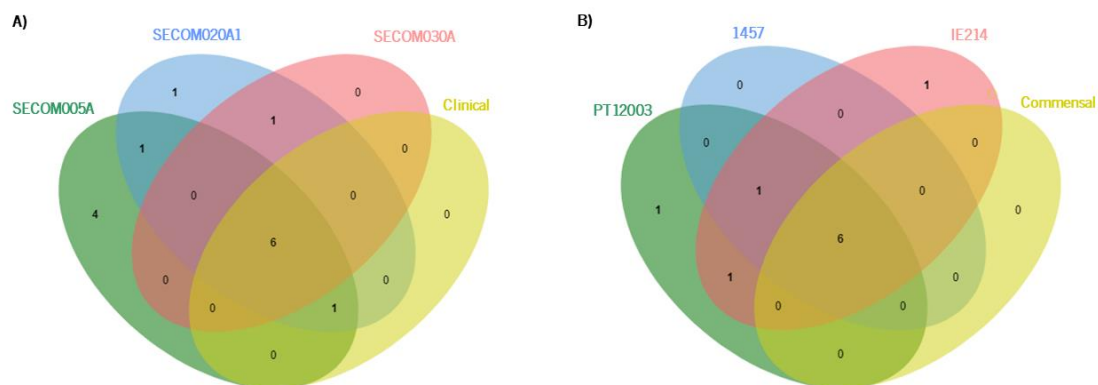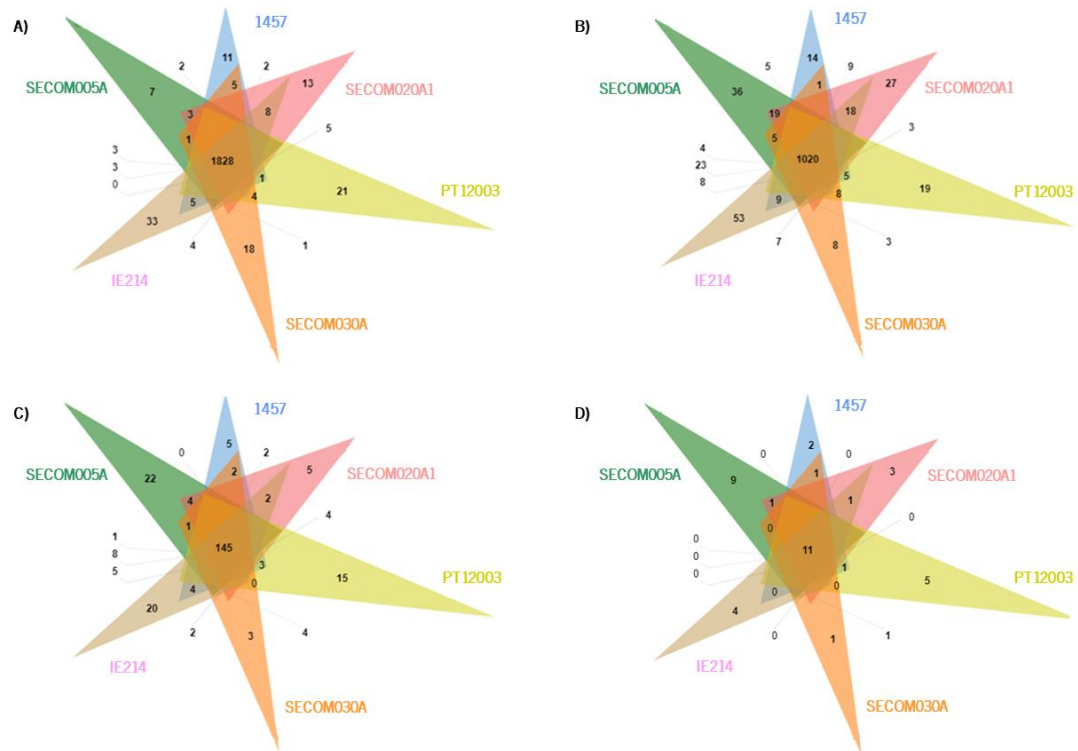
Figure 4.4. Venn diagrams showing the number of common transcripts (overlapping circles) and unique transcripts (non -overlapping circles) between commensal and clinical isolates with different RPKM values: RPKM >10 (A); RPKM >100 (B); RPKM >1000 (C); RPKM > 5000 (D).

By analyzing commonly expressed genes, using different RPKM thresholds, we found 1951 common genes at RPKM >1, 1828 at RPKM >10, 1020 at RPKM >100, 145 at RPKM >1000 and 11 at RPMK >5000. Then, the genes that were present in the lower RPKM threshold but absent in the next tier of RPKM threshold were selected. A total of 1944 genes were identified, and their discriminative power was analyzed. However, it should be noted that, contrary to the first strategy used, which highlights genes only detected in one of the groups, in this third analysis, the highlighted genes might not be good candidates, because the detected difference in the expression levels might not relate to be a true discriminative power, as exemplified in Figure 4.5.

Table 4.5. RPKM values of the genes only detected in clinical or commensal isolates through the third strategy of analysis

Genes	Clinical isolates			Commensal isolates		
	1457	IE214	PT12003	SECOM005A	SECOM020A1	SECOM030A
SERP_RS10265	41.53	34.12	18.77	6.09	8.99	4.49
SERP_RS04470	28.78	43.78	85.31	1.91	8.77	5.63
SERP_RS11115	10.96	19.90	12.94	4.58	4.01	3.00
SERP_RS00240	10.56	12.78	11.33	<1	9.65	7.23
SERP_RS10580	100.1	347.5	116.0	82.60	53.83	82.90
SERP_RS02550	100.2	108.3	101.3	81.20	53.30	87.81
SERP_RS08685	106.0	131.1	105.0	51.95	70.30	84.6
SERP_RS00880	120.0	137.9	114.4	66.56	86.22	92.3
SERP_RS12665	177.0	371.3	137.2	<1	<1	80.86
SERP_RS05295	275.3	132.8	426.3	79.79	79.37	68.35
SERP_RS13425	<1	6.07	8.97	34.94	122.3	22.91
SERP_RS03740	<1	<1	<1	5518.4	149.4	30.55
SERP_RS11770	6.56	<1	7.04	36.54	11.99	35.94
SERP_RS11970	<1	<1	<1	148.8	170.3	81.75
SERP_RS09755	63.50	1.47	92.80	729.4	216.8	176.8
SERP_RS10630	67.80	99.30	85.59	117.3	105.8	113.8
SERP_RS11445	63.80	59.90	89.70	128.7	107.6	113.0
SERP_RS05705	872.5	995.0	601.1	1936.5	1402.7	1075.2

Not surprisingly, the theoretical differential expression of most of the genes analyzed was not discriminative between both groups. Of all the genes included in this third analysis, only eleven were found to potentially be able to discriminate clinical from commensal isolates, including genes previously highlighted in the first analysis. The function or putative function of these genes is shown in Table 4.6.
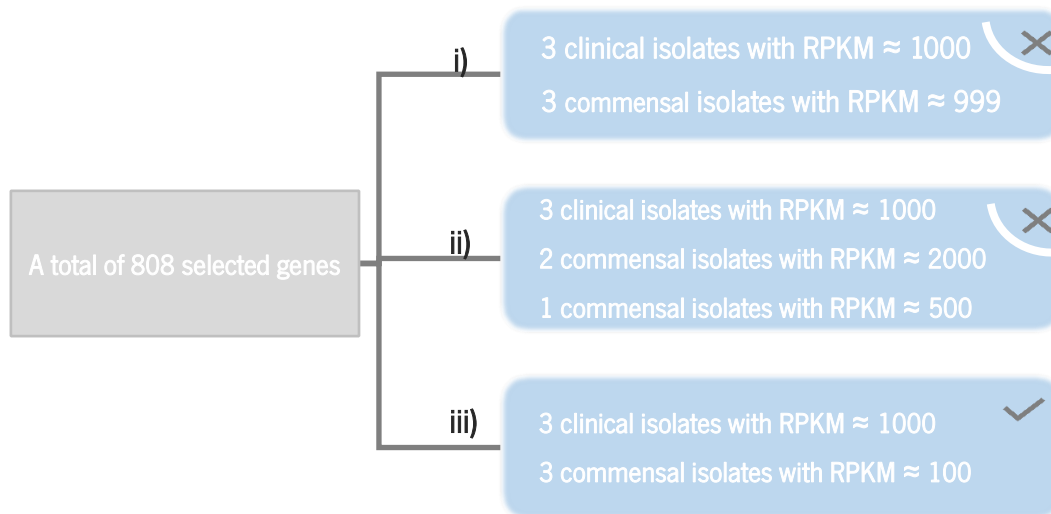
Figure 4.5. Differential RPKM threshold analysis require validation of the highlighted genes. Some examples of situations that requires validation in a scenario of RPKM threshold of 1000: **(i)** no significant differences in the expression of the genes in the 6 isolates, but the imposed RPKM threshold of 1000 wrongly highlights this as a potential maker; **(ii)** if just one of the isolates has a differential level of expression, the gene will be wrongly highlighted as a potential discriminative gene; **(iii)** the ideal situation occurs when the genes are highlighted under a specific RPKM threshold and the theoretical discriminative power is sufficient enough.

4.3.2 RNA-seq validation

To confirm the theoretical discriminative power of the genes highlighted before, qPCR was used to first validate the RNA-seq data (technical validation) and then to confirm those observations in independent experiments (biological validation). This was absolutely necessary because RNA-seq data were obtained from only one single run, which nevertheless included a pool of independent RNA obtained from multiple experiments. qPCR is considered the gold standard for measuring gene expression and it is normally used to validate the results obtained by RNA-seq since a high correlation is observed between these two methodologies [7].

.

Table 4.6. List of the genes of interest, with fold-change values (based on RPKM values)

Gene	Definition	Clinical Isolates (RPKM)	Commensal isolates (RPKM)	Fold-change (Clinical/Commensal)
		Mean ± SD	Mean ± SD	
SERP_RS00240	Cro/CI family transcriptional regulator	11.6 ± 1.13	5.63 ± 5.02	2.05
SERP_RS10265	PAP2 family protein	31.5 ±11.6	6.52 ± 2.28	4.8*
SERP_RS11115	DUF1672 domain-containing protein	14.6 ± 4.70	3.86 ± 0.80	3.78*
SERP_RS04465	ABC transporter ATP-binding protein	48.5 ± 35.3	7.18 ± 3.47	6.76
SERP_RS04470	ABC transporter permease	52 ± 29.2	5.44 ± 3.43	9.69*
SERP_RS04475	Sensor histidine kinase	39.1 ± 10.9	8.70 ± 2.46	4.49**
SERP_RS05305	Oxidoreductase short-chain dehydrogenase family	65.3 ± 13.2	23.8 ± 9.59	2.7*
SERP_RS06815	Uncharacterized protein	743.4 ± 163.8	347.9 ± 32.8	2.13*
SERP_RS11970	MFS transporter	0.27 ± 0.37	133.6 ± 46.2	-504**
SERP_RS11770	Phenol soluble modulin beta 1	4.53 ± 3.93	28.2 ± 14.0	-6.2*
SERP_RS03740	Phenol soluble modulin beta 1	2.48 ± 4.29	1899 ± 3135	-726
SERP_RS10985	Universal stress protein	9.58 ± 2.44	82.2 ± 27.1	-8.6**

Data represented the mean ± standard deviation of RPKM values of all isolates. Statistical significance was analyzed using unpaired *t* test * $p < 0.05$, * *$p < 0.01$

For the technical validation, aliquots of the total RNA used for the RNA-seq analysis were used. As shown in Figure 4.6, the same trend (up or downregulation) obtained in RNA-seq data was observed in most of the tested genes, with only one exception, the gene *SERP_RS00240*. Furthermore, some significant differences were found in the absolute differential expression, wherein *SERP_RS05305* and *SERP_RS04465* genes had significant lower discriminative potential when quantified by qPCR as in comparison with the RNA-seq data.
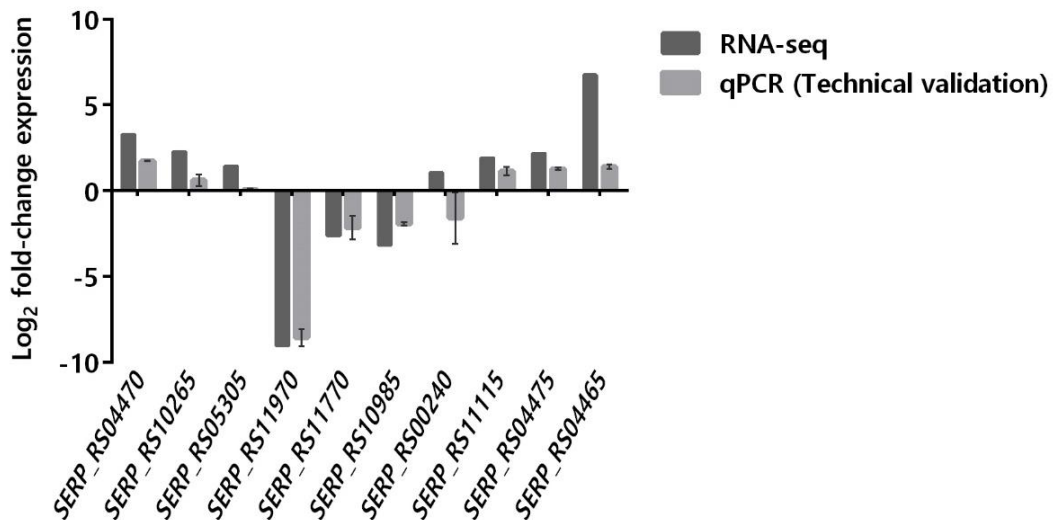
Figure 4.6. Technical validation, by qPCR, of the level of transcription of the selected genes. The data indicate the fold-change expression of genes between clinical isolates and commensal isolates. The bars represent the mean ± SEM of three technical replicates.

Although qPCR and RNA-seq are both used to measure gene expression, the normalization is performed differently. In qPCR reference genes are used for normalization and the measurement of expression is relative to the expression of reference gene [30] while RNA-seq measures absolute expression levels [31]. In this study, it was observed that the fold differences in expression when using RNA-seq were higher than the ones obtained using qPCR (an example is given in Table 4.7). These differences might occur due to artifacts created during library construction, which may affect the comparison across genes [7]. Furthermore, the majority of the selected genes were expressed at low levels and it is known that low expressing transcripts are more variable [32, 33]. Thus, under these conditions, the correlation between qPCR and RNA-seq is lower [34]. Importantly, while the average of values obtained for a condition might be statistically significant when compared with others, often one of the three isolates within a group

has a variable expression that reassembles more the three isolates from the other group, which is a bad indicator for future discriminative potential. An example is given in Table 4.7.

Table 4.7. Analysis of expression of *SERP_RS04470* gene obtained by RNA-seq and qPCR and respective fold-change values (clinical/commensal)

	Clinical isolates			Commensal isolates			Fold-change
	1457	IE214	PT12003	SECOM005A	SECOM020A1	SECOM030A	
RPKM	28.78	43.78	85.31	1.91	8.77	5.63	9.69*
qPCR	9.63E-6	2.17E-5	4.04E-6	1.74E-6	4.15E-6	4.76E-6	3.32

Statistical significance was analyzed using unpaired t test. *$p < 0.05$

Having this drawback in consideration, only two of the tested genes, *SERP_RS11970* and *SERP_RS10985*, had sufficient discriminative power not only between the averages of both groups but between the six strains tested. These genes codify a transporter (*SERP_RS11970*) and a universal stress protein (*SERP_RS10985*). Because there is a lot of variation in bacterial gene expression when incubated in human blood [35, 36] the two most promising genes were further tested, by performing biologic validation, by repeating the same experiment but using different blood donors. It is well known that there is a significant source of variability inherent to each donor such as age [37], gender [38] and the proportion of different cell populations [39] as described previously in Chapter 3. Moreover, the genetic background of the host seems to affect the transcriptional response of pathogens [40]. Therefore, as shown in Figure 4.7, while *SERP_RS10985* maintained the expression stable, in accordance with RNA-seq data, a significant decrease in the discriminative power of *SERP_RS11970* was observed. Nevertheless, the differences in the transcription levels between technical and biological validations did not exclude *SERP_RS11970* as a gene with the potential to discriminate commensal from clinical isolates.
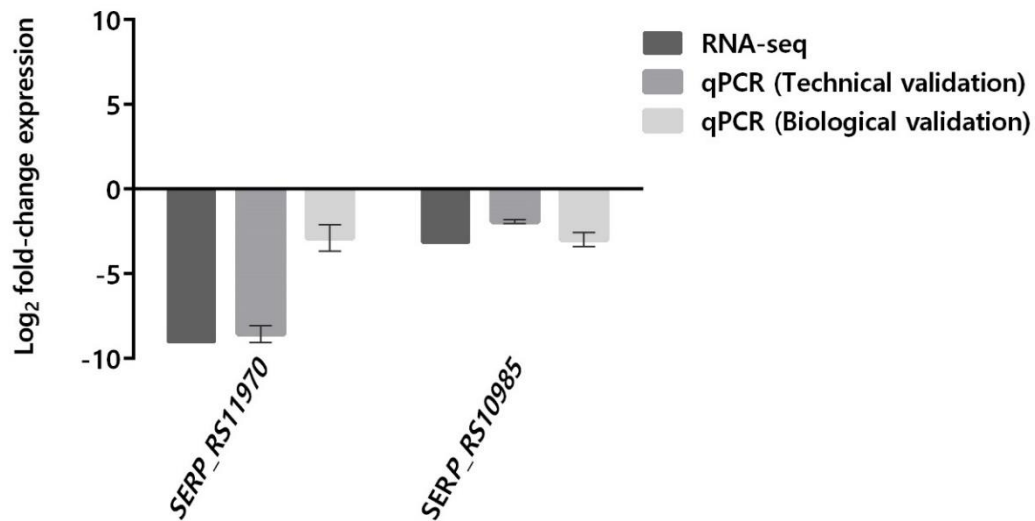
Figure 4.7. Validation, by qPCR, of the expression levels of the genes *SERP_RS11970* and *SERP_RS10985*.
The data indicate the fold-change between clinical and commensal isolates. For technical validation, the bars represent the mean ± SEM of three technical replicates and for biologic validation, the bars represent the mean ± standard deviation of three independent experiments, using three different donors

4.4 CONCLUSION

In conclusion, the comparison of the six transcriptomes after two hours of incubation with human blood revealed that the gene *SERP_RS11970* was the best potential molecular marker to discriminate clinical from commensal isolates. However, since it is known that gene expression is highly variable [36], the transcription levels of the most promising genes, highlighted in this analysis, will need to be confirmed under other conditions, namely after four and six hours of incubation in human blood. Furthermore, the validation of the discriminative potential of the molecular marker selected will need to be also addressed in a representative worldwide collection of *S. epidermidis* isolates. This will be explored in Chapter 5.

4.5 REFERENCES

1. Kleinschmidt S, Huygens F, Faoagali J, Rathnayake IU, Hafner LM. *Staphylococcus epidermidis* as a cause of bacteremia. *Future Microbiology* 2015;10(11):1859-1879.

2. Conlan S, Mijares LA, Becker J, Blakesley RW, Bouffard GG *et al*. *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biology* 2012;13(7):2012-2013.

3. Kleinschmidt SL. *Molecular characterisation of Bacteraemia and contamination isolates of Staphylococcus epidermidis isolated from patient blood cultures*. Queensland University of Technology; 2018.

4. Pinheiro L, Brito CI, Oliveira A, Pereira VC, Cunha Mde L. *Staphylococcus epidermidis* and *Staphylococcus haemolyticus*: detection of biofilm genes and biofilm formation in blood culture isolates from patients in a Brazilian teaching hospital. *Diagnostic microbiology and infectious disease* 2016;86(1):11-14.

5. Hellmark B, Soderquist B, Unemo M, Nilsdotter-Augustinsson A. Comparison of *Staphylococcus epidermidis* isolated from prosthetic joint infections and commensal isolates in regard to antibiotic susceptibility, agr type, biofilm production, and epidemiology. *International Journal of Medical Microbiology* 2013;303(1):32-39.

6. Cheung AL, Bayer AS, Zhang G, Gresham H, Xiong YQ. Regulation of virulence determinants *in vitro* and *in vivo* in *Staphylococcus aureus*. *FEMS Immunology Medical Microbiology*, Review 2004;40(1):1-9.

7. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, Review 2009;10(1):57-63.

8. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A *et al*. A survey of best practices for RNA-seq data analysis. *Genome Biology* 2016;17(1):13.

9. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 2008;5(7):621-628.

10. Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E *et al*. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics* 2009;25(24):3207-3212.

11. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research* 2008;18(9):1509-1517.

12. Croucher NJ, Thomson NR. Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology* 2010;13(5):619-624.

13. França A, Pier GB, Vilanova M, Cerca N. Transcriptomic Analysis of *Staphylococcus epidermidis* Biofilm-Released Cells upon Interaction with Human Blood Circulating Immune Cells and Soluble Factors. *Frontiers in Microbiology*.

14. Franca A, Carvalhais V, Vilanova M, Pier GB, Cerca N. Characterization of an *in vitro* fed-batch model to obtain cells released from *S. epidermidis* biofilms. *AMB Express* 2016;6(1):016-0197.

15. Cerca N, Jefferson KK, Oliveira R, Pier GB, Azeredo J. Comparative Antibody-Mediated Phagocytosis of *Staphylococcus epidermidis* Cells Grown in a Biofilm or in the Planktonic State. *Infection and Immunity* 2006;74(8):4849-4855.

16. Freitas AI, Lopes N, Oliveira F, Bras S, Franca A *et al.* Comparative analysis between biofilm formation and gene expression in *Staphylococcus epidermidis* isolates. *Future Microbiology* 2018;13:415-427.

17. Cerca N, Pier GB, Vilanova M, Oliveira R, Azeredo J. Quantitative analysis of adhesion and biofilm formation on hydrophilic and hydrophobic surfaces of clinical isolates of *Staphylococcus epidermidis*. *Research in Microbiology* 2005;156(4):506-514.

18. Oliveira F, Cerca N. Antibiotic resistance and biofilm formation ability among coagulase-negative staphylococci in healthy individuals from Portugal. *Journal of Antibiotics* 2013;66(12):739-741.

19. Oliveros JC. An interactive tool for comparing lists with Venn diagram version 2.1.0. 2007.

20. Hebenstreit D, Fang M, Gu M, Charoensawan V, van Oudenaarden A *et al.* RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Molecular System Biology*, Comparative Study 2011;7(497):28.

21. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* 2013;41:29.

22. Pundir S, Martin MJ, O'Donovan C. UniProt Tools. *Current Protocols Bioinformatics* 2016;53(1):1-15.

23. Sousa C, Franca A, Cerca N. Assessing and reducing sources of gene expression variability in *Staphylococcus epidermidis* biofilms. *Biotechniques* 2014;57(6):295-301.

24. Freitas AI, Vasconcelos C, Vilanova M, Cerca N. Optimization of an automatic counting system for the quantification of *Staphylococcus epidermidis* cells in biofilms. *Journal Basic Microbiology* 2014;54(7):750-757.

25. França A, Freitas AI, Henriques AF, Cerca N. Optimizing a qPCR Gene Expression Quantification Assay for *S. epidermidis* Biofilms: A Comparison between Commercial Kits and a Customized Protocol. *PLOS ONE* 2012;7(5):e37480.

26. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC *et al.* Primer3–new capabilities and interfaces. *Nucleic Acids Research* 2012;40(15):22.

27. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25(4):402-408.

28. Pfaffl MW. Quantification strategies in real time PCR. *A-Z of quantitative PCR*. La Jolla, CA, USA: International University Line; 2004. pp. 87-112.

29. Fang Z, Cui X. Design and validation issues in RNA-seq experiments. *Brief Bioinformatics*, Review 2011;12(3):280-287.

30. Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal Biomolecular Technology* 2004;15(3):155-166.

31. t Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH *et al.* Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Research* 2008;36(21):15.

32. Raz T, Kapranov P, Lipson D, Letovsky S, Milos PM *et al.* Protocol dependence of sequencing-based gene expression measurements. *PLOS ONE* 2011;6(5):0019287.

33. Ramskold D, Kavak E, Sandberg R. How to analyze gene expression using RNA-sequencing data. *Methods Molecular Biology* 2012;802:259-274.

34. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 2008;320(5881):1344-1349.

35. Vandecasteele SJ, Peetermans WE, Merckx R, Van Eldere J. Expression of biofilm-associated genes in *Staphylococcus epidermidis* during in vitro and in vivo foreign body infections. *Journal Infection Disease* 2003;188(5):730-737.

36. França A, Carvalhais V, Maira-Litrán T, Vilanova M, Cerca N *et al.* Alterations in the *Staphylococcus epidermidis* biofilm transcriptome following interaction with whole human blood. *Pathogens and Disease* 2014;70(3):444-448.

37. Eady JJ, Wortley GM, Wormstone YM, Hughes JC, Astley SB *et al.* Variation in gene expression profiles of peripheral blood mononuclear cells from healthy volunteers. *Physiol Genomics* 2005;22(3):402-411.

38. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC *et al.* Individuality and variation in gene expression patterns in human blood. *Proceedings of the National Academy of Sciences* 2003;100(4):1896-1901.

39. **Cobb JP, Mindrinos MN, Miller-Graziano C, Calvano SE, Baker HV** *et al.* Application of genome-wide expression analysis to human health and disease. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(13):4801-4806.

40. **Thänert R, Goldmann O, Beineke A, Medina E**. Host-inherent variability influences the transcriptional response of *Staphylococcus aureus* during in vivo infection. *Nature communications*, Article 2017;8:14268.

CHAPTER 5

Validation of potential diagnostic target using a diverse collection of clinical and commensal *S. epidermidis* isolates

SUMMARY

The discrimination between *S. epidermidis* isolates that cause bloodstream infections from contamination is a daily challenge for clinicians. The identification of genes differentially expressed between *S. epidermidis* clinical and commensal isolates, upon contact with human blood, could be the key to identify differences in the ability to cause infection. In the previous chapter, markers with the potential to differentiate between three clinical isolates from three commensal isolates were found. Herein, those genes were further tested to determine if they could differentiate commensal from clinical isolates, using bigger control groups. The comparison of the transcription levels within both groups revealed that the differences previously found were no longer statistically different when using 50 isolates. This result reinforced the hypothesis that *S. epidermidis* is an opportunistic pathogen. To further validate that hypothesis, qPCR data were also compared with bacterial survivability in human blood. No significant differences were observed between both groups.

5.1 BRIEF INTRODUCTION

As mentioned previously in Chapter 2, the ubiquitous nature of *S. epidermidis* infections strongly contributes to the difficulties in making a correct diagnosis [1]. Clearly identifying differentiation traits between clinical (invasive) and commensal *S. epidermidis* isolates is expected to help the diagnostic process. In the previous chapter, a transcriptomic approach was used in an attempt to identify potential discriminatory markers, by evaluating bacterial gene expression when co-cultured in human blood.

As the previous analysis was performed on a limited number of isolates, due to experimental costs, herein, the most promising genes identified before were further validated using a bacterial collection from different sources, such as 26 clinical isolates, 24 commensal isolates from the community and also 6 more from hospital colonization and contamination. Furthermore, a phenotypical analysis was also performed and compared with the gene expression data, by determining the survivability of the isolates upon co-culture in human blood.

5.2 MATERIAL AND METHODS

5.2.1 Bacterial Strains

S. epidermidis nosocomial isolates used in this study are listed in Table 5.1. Isolates from different countries were considered as being from infection or contamination, according to local clinicians' evaluation criteria. *S. epidermidis* commensal isolates from the community are listed in Table 5.2.

5.2.2 Growth conditions

One single colony was inoculated into 3 mL of Tryptic Soy Broth (TSB) (Liofilchem, Teramo, Italy) and incubated for 16h, at 37$\circ$C and 120 rpm (ES-20 Shaker-Incubator, Biosan, Riga, Latvia). Thereafter, cells were harvested by centrifugation at 16000 g for 7m at 4$\circ$C, washed twice and resuspended in 0.9% NaCl, in order to obtain a final concentration of approximately 10^9 CFU/mL. These suspensions were used for subsequent co-incubation with human or defibrinated horse blood.

Table 5.1. List of *S. epidermidis* nosocomial isolates used in this study

Strain	Country	Year	Gender	Age	Origin	Source	Reference
					Nosocomial isolates		
RP62A	USA	1979	-	-	Infection	Blood	[2]
1457	Germany	1989	-	-	Infection	Central Venous Catheter	[3]
IE186	USA	-	-	-	Infection	Endocarditis	[4]
IE214	USA	-	-	-	Infection	Endocarditis	
COB17	Colombia	1997	-	-	-	Urine	
COB20	Colombia	1997	-	<1	Infection	Urine	
DEN19	Denmark	1997	M	76	Infection	-	
DEN120	Denmark	1998	M	27	Infection	-	
GRE26	Greece	1998	-	-	-	-	
HUR51	Hungary	1997	M	44	Infection	Wound	
ICE102	Iceland	1998	M	76	Infection	Wound	
URU23	Uruguay	1997	M	37	Infection	Urine	
ICE21	Iceland	1997	F	57	Infection	Wound	[5]
ICE24	Iceland	1997	M	76	-	Wound	
ICE5	Argentina	1997+	M	79	Infection	Urine	
ICE9	Iceland	1997	F	68	Infection	Urine	
ITL34	Italy	1997	M	41	Infection	Wound	
MCO150	Mexico	1998	M	6	Infection	Blood	
MEX37	Mexico	1996	-	-	Infection	Cerebrospinal fluid	
MEX60	Mexico	1996	-	<12	Infection	Catheter	
PLN64	Poland	1997	M	50	Infection	Wound	
TW113	Taiwan	1997	M	76	Infection	Respiratory	

69

Table 5.1. Continued

Strain	Country	Year	Gender	Age	Origin	Source	Reference
ESP43	Spain	1997	F	59	Infection	Blood	[5]
PT11002	Portugal	2011	M	5	Infection	Blood	
PT11015	Portugal	2011	F	25	Infection	Blood	
PT12003	Portugal	2011	M	62	Surgical site Infection	Central catheter	
PT12005	Portugal	2012	M	75	Surgical site Infection	Blood	
PT12008	Portugal	2012	F	65	Catheter infection	Expectoration	
PT12013	Portugal	2012	M	90	Respiratory tract Infection	Blood	
PT12023	Portugal	2012	M	69	Infection	Blood	[6]
PT12030	Portugal	2012	M	<1	Infection	Blood	
PT12032	Portugal	2012	F	66	Infection	Blood	
PT12050	Portugal	2012	M	56	Infection	Blood	
PT12060	Portugal	2012	M	50	Infection	Blood	
PT12065	Portugal	2012	M	75	Infection	Blood	
PT13011	Portugal	2012	-		Infection	Blood	
PT13038	Portugal	2012	F	75	Infection	Blood	
PT13042	Portugal	2013	M	43	Infection	Blood	
DEN116	Denmark	1998	M	71	Contamination	Blood	
DEN185	Denmark	1998	M	41	Contamination	Blood	
DEN69	Denmark	1997	F	35	Contamination	Blood	[5]
ICE192	Iceland	1998	M	2	Contamination	Blood	
DEN94	Denmark	1997	-	79	Contamination	Blood	
PE9	Boston	-	-	-	Colonization	-	[7]
CV45	Cape Verde	1997	F	10	Colonization	-	
DEN110	Denmark	1998	M	50	Colonization	-	[5]

70

Table 5.2. List of *S. epidermidis* commensal isolates from community used in this study

	Commensal isolates from community					
Strain	Country	Year	Gender	Age	Origin	Reference
SECOM005A	Portugal	2012	F	23		
SECOM20A1	Portugal	2012	M	15		
SECOM030A	Portugal	2012	F	45		
SECOM001B	Portugal	2012	F	56		
SECOM003A	Portugal	2012	M	35		
SECOM010B	Portugal	2012	M	52		
SECOM022A	Portugal	2012	F	14		
SECOM023A	Portugal	2013	M	22		
SECOM024A	Portugal	2013	M	17		
SECOM027A	Portugal	2013	F	48		
SECOM029A	Portugal	2013	M	50		
SECOM031A	Portugal	2013	M	25		
SECOM034A	Portugal	2013	M	19	Community isolates from healthy individuals	[8]
SECOM035A	Portugal	2013	M	20		
SECOM037A	Portugal	2013	M	19		
SECOM040A	Portugal	2013	F	19		
SECOM042A	Portugal	2013	F	20		
SECOM049A	Portugal	2013	M	28		
SECOM053A	Portugal	2013	M	27		
SECOM058A	Portugal	2013	M	48		
SECOM062A	Portugal	2013	M	10		
SECOM066A	Portugal	2013	F	19		
SECOMF12	Portugal	2013	F	21		
SECOMM14	Portugal	2013	M	21		

5.2.3 Gene detection by PCR

In order to obtain genomic DNA, one to five colonies of each isolate were collected from a Tryptic Soy Agar (TSA), which was prepared using TSB (VWR, Leuven, Belgium) supplemented with 1.5% agar (Liofilchem, Teramo, Italy) and inoculated into 200 µL of nuclease-free water. The suspension was incubated in a heating block at 95∘C for 10 min and immediately cooled in ice for 5 min in order to disrupt cell wall and release the cell content. The lysate was then centrifuged at 13500 g during 5 min. One µL of the supernatant was used as template for PCR amplification in 10 µL of reaction volume and containing 5 µL of NZYtaq II Green Master mix (NZYtech, Lisboa, Portugal), 1 µL of primer mixture with a 0.5 µM concentration each, and 3 µL of nuclease-free water. The primers used are listed in Chapter 4. The PCR was performed in the MJ Mini thermal cycler (Bio-Rad, Hercules, CA) with the following cycling parameters: 5 min at 95∘C followed by 35 repeats of 30s at 95∘C, 15s at 58∘C or 60∘C and 45s at 72∘C. PCR products were analyzed by gel electrophoresis with 1% agarose (Bio-Rad, Hercules, CA) stained with Midori Green DNA stain (Nippon Genetics Europe GmbH) and visualized by ChemiDoc™ XRS+ (Bio-Rad, Hercules, CA). A 100-bp DNA ladder (NZYTech, Lisboa, Portugal) was used as a marker.

5.2.4 Human blood collection

Peripheral blood was collected from healthy adult volunteers, not taking antibiotics or anti-inflammatory medication within the last 14 days, by venipuncture into BD Vacutainer® tubes spray coated with lithium heparin (Becton Dickinson, NJ, USA). Blood was collected under a protocol approved by the Institutional Review Board of the University of Minho (SECVS 002/2014 (ADENDA)), which is in strict accordance with the Declaration of Helsinki and Oviedo Convention. All donors gave written informed consent to have blood taken.

5.2.5 Co-incubation of bacteria with human and horse blood

For co-incubation, two different experimental set-up were performed. For analysis of gene expression at different time points, the co-incubation was performed using IE214 and SECOM020A isolates and two different types of blood: heparinized human blood and defibrinated horse blood (Thermo Fisher Scientific, MA, USA). Briefly, in 2 mL tubes, 50 µL of a suspension of 1×10^9 CFU/mL were mixed with 450 µL of blood and incubated at 80 rpm (PSU-10i) for 2, 4 and 6h at 37°C. This assay was performed two to three independent times. For the analysis of gene expression using a representative worldwide collection of *S. epidermidis* isolates, the co-incubation was performed during 4h in defibrinated horse blood (Thermo Fisher Scientific, MA, USA) using the same procedure. After the co-

incubation period, in order to lyse eukaryotic cells, samples were sonicated for 5s at 33% amplitude and, then, bacteria were harvested by 5 min centrifugation at 16000 g at 4 °C.

5.2.6 *S. epidermidis* gene expression assays

RNA extraction, cDNA synthesis and qPCR were performed as described in Chapter 4.

5.2.7 Bacterial survival

The bacterial suspensions were prepared as described above and by adjusting the concentration to 10^9 CFU/mL. After that, serial dilutions (10-fold) were performed in order to obtain a concentration of 10^6 CFU/mL. Then, 50 µL of these suspensions were mixed with 450 µL of human blood and, subsequently, incubated at 37 °C, at 80 rpm. The number of surviving bacteria was assessed after 4h of incubation with human blood. After incubation, an aliquot was taken, and the enumeration of bacteria was performed by CFU counting. The number of CFU/mL before incubation was used as control to calculate the percentage of survival. This experiment was performed two to three independent times, using blood from different donors.

5.2.8 Statistical analysis

Statistical analysis was carried out with GraphPad Prism Version 6 Trial (CA, USA). For comparisons among different groups, one-way or two-away ANOVA, with Tukey´s or Bonferroni´s multiple comparisons tests, were used when appropriate (the tests used are detailed in the figure caption). *P* < 0.05 was considered significant.

5.3 RESULTS AND DISCUSSION

5.3.1 Gene expression at different time points using one donor or different donors

Taken into consideration the results obtained in the previous chapters, we first evaluated the expression of the gene *SERP_RS11970,* up to 6h of incubation in human blood. This was necessary since RNA-seq was performed only after 2h of incubation and it is known that gene expression is highly variable [9]. Furthermore, it was previously shown that in the first few hours upon medium change, which occurs in fed-batch systems, the bacterium's transcriptome changes in a non-specific manner [10]. This gene was the most promising differentially expressed in commensal isolates compared to clinical isolates. Therefore, one clinical (IE214) and one commensal isolate (SECOM020A1), that showed a pronounced difference in expression levels of this gene during the experiments described in Chapter 4, were used for these experiments. Figure 5.1 demonstrates that no significant differences

were found between different time points for both genes, suggesting that the niche adaptation provided by the products of this gene occurs rapidly upon contact with human blood and remains constant, at least during the 6h tested.
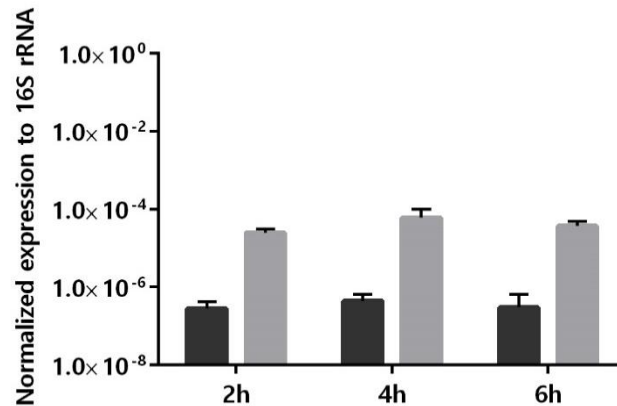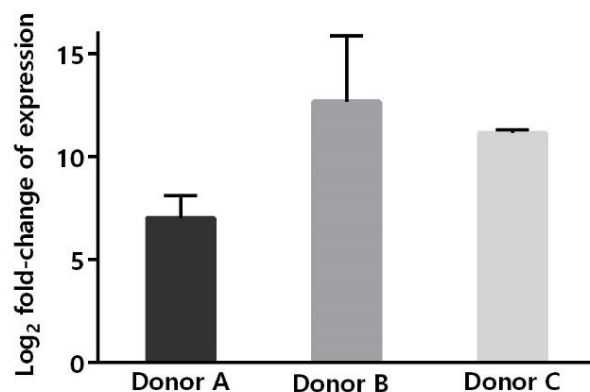

Figure 5.1. Quantification of the transcription levels of *SERP_RS11970* gene after 2, 4 and 6h of incubation in human blood using the same donor. The data indicate the normalized expression. The bars represent the mean plus standard deviation of three independent experiments using the same donor. Statistical differences between groups were analyzed with two-way ANOVA and Tukey´s multiple comparison test. $p > 0.2090$.

Next, we assessed the possible variability inherent to different donors. For this experiment, the expression of the target gene was assessed after 4h of incubation. This time point was selected as a compromise between stable gene expression and time needed to perform the experiments. Furthermore, as shown before, at this time point a lower number of unspecific regulation of genes occurs [10]. Interestingly, while significant differences were found between the different donors, the overall trend was maintained (Figure 5.2), with *SERP_RS11970* being, on average, 10.28-fold more expressed in commensal than in clinical isolates.

5.3.2 Gene expression using fresh human blood *vs* defibrinated horse blood

One recurrent problem with the current experimental design was the low availability of human blood donors. Since it was previously demonstrated that the human blood soluble factors, and not the cell fraction, strongly contribute to the alterations in gene expression in *S. epidermidis* [11], we hypothesized that the utilization of commercially available defibrinated horse blood, which is deprived of leukocytes [13], could be a viable alternative to further proceed with the subsequent experimental goals (the biological validation using a large number of strains).

Figure 5.2. Quantification of the transcription levels of *SERP_RS11970* gene after 4h of incubation in human blood using different donors. The data indicate the fold-change of expression of *SERP_RS11970* gene between commensal and clinical isolates. The bars represent the mean plus standard deviation of three technical replicates. Statistical differences between groups were analyzed with one-way ANOVA and Tukey´s multiple comparison test. $p > 0.05$.

Therefore, in order to reduce the utilization of human blood, commercially available defibrinated horse blood was used to determine the expression levels of *SERP_RS11970*. As shown in Figure 5.3, the same trend was observed when the bacterial cultures were incubated in either human blood or defibrinated horse blood, confirming that *SERP_RS11970* expression is equally induced in both types of blood. Interestingly, it was also observed that the fold-change between commensal and clinical isolates in horse blood was higher than obtained in human blood, except after 6h of incubation. Besides that, as expected, it was also noticed that in human blood, an increase in variability was observed, due to the fact that different donors were used in different experiments.

To further confirm that horse blood could be used as a substitute of human blood, in the scope of this experimental design, we repeated the co-culture experiments, using two more clinical isolates (DEN120 and ICE5) and two more commensal isolates (SECOM027A and SECOM029A). Also, other four genes, highlighted in the previous chapter, were also included in this study. These genes (*SERP_RS10265, SERP_RS05305, SERP_RS06815*, and *SERP_RS04470)* were selected based on RPKM values obtained after 2h of incubation with human blood and based on their function.
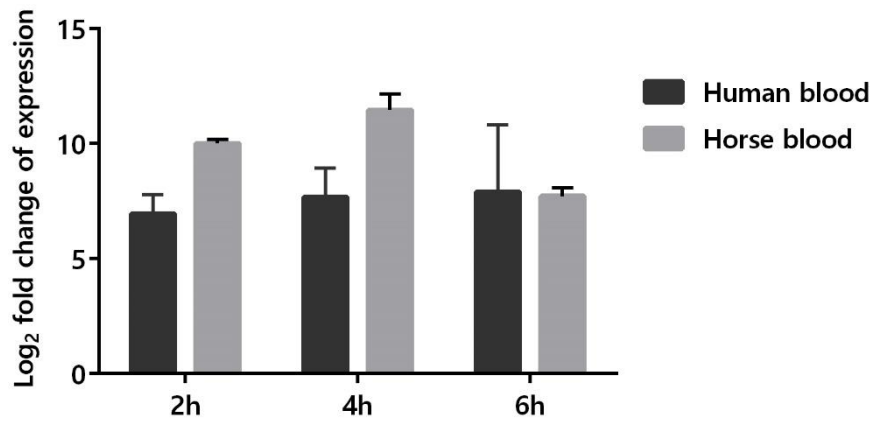
Figure 5.3. Quantification of the transcription levels of *SERP_RS11970* gene after 2, 4 and 6h of incubation in human blood (two different donors) and horse blood. The data indicate the fold-change of expression between commensal and clinical isolates. The bars represent the mean plus standard deviation of two independent experiments. Statistical differences between groups were analyzed with two-way ANOVA and Bonferroni´s multiple comparison test. $p > 0.2090$.

5.3.3 Gene expression using a worldwide collection of *S. epidermidis* isolates

It is well known that the population of *S. epidermidis* presents a high level of diversity in isolates from different geographical or clinical origins, but also within collections originated in the same hospital [5]. To enclose this diversity of isolates, the transcription levels of the selected genes were analyzed using a representative and diverse worldwide *S. epidermidis* isolates. This collection included 38 isolates from infection and 24 isolated from the skin of healthy individuals with no contact with the hospital environment. First, the presence of the genes of interest, selected in the experiments described in Chapter 4, was assessed by PCR in all 62 isolates.

Figure 5.4 shows the normalized expression of the selected genes after 4h of incubation in human or horse blood. In general, it can be observed that the transcription levels of the selected genes were similar when co-incubation occurred in either human or horse blood. Thus, it appears that the selected genes are mainly mediated by plasma components since transcription levels were not altered by the use of horse blood.

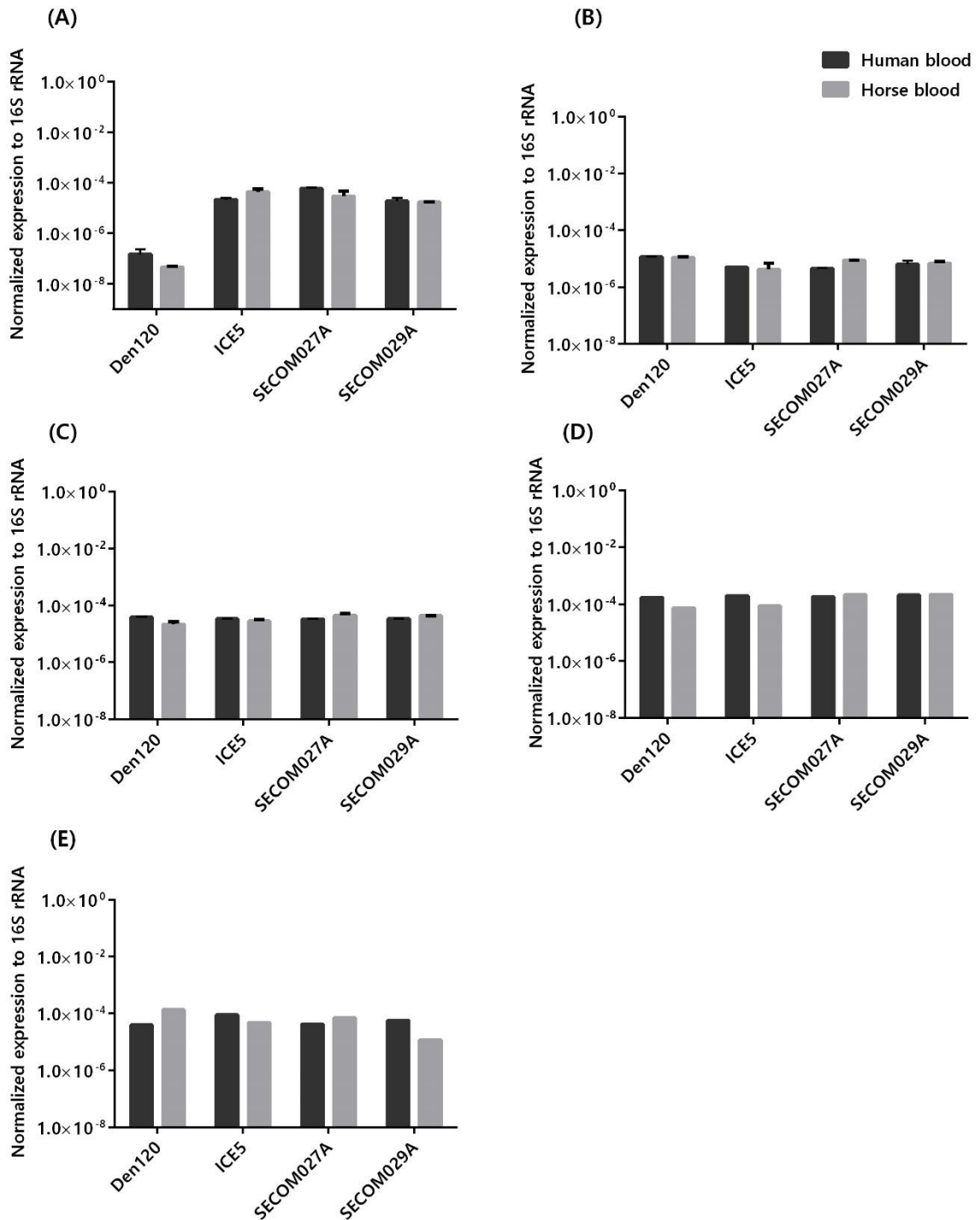Figure 5.4. Comparison of transcriptions levels of the selected genes after 4h of incubation in human or horse blood. (A) *SERP_RS11970* (B) *SERP_ RS10265* (C) *SERP_RS05305* (D) *SERP_RS06815* and, (E) *SERP_RS04470.* The bars represent the mean plus standard deviation of two technical replicates. Statistical differences between groups were analyzed with one-way ANOVA and Bonferroni´s multiple comparison test. $p > 0.05$.

As shown in Table 5.4, the *SERP_RS11970* gene was only detected in 26 out of 38 isolates from infection while in commensal isolates it was found in all 24 tested isolates. Strikingly, all other genes were present in all tested isolates. However, because *SERP_RS11970* had the most promising discriminative power, the follow-up studies included only the 50 isolates that harboured this gene.

Table 5.4. Prevalence of genes in a worldwide collection of *S. epidermidis* isolates

Genes	Clinical isolates (n=38)		Commensal isolates (n=24)	
	Number of isolates PCR positive	%	Number of isolates PCR positive	%
SERP_RS11970	26	68	24	100
SERP_RS10265	38	100	24	100
SERP_RS04470	38	100	24	100
SERP_RS05305	38	100	24	100
SERP_RS06815	38	100	24	100

Unfortunately, as shown in Figure 5.5A, the discriminative power of the expression of the gene *SERP_RS11970* was lost, once a wider collection of *S. epidermidis* isolates was used. Apparently, the transcription levels of *SERP_RS11970* gene within the group of clinical isolates presented extraordinary levels of variability, with some strains expressing this gene more than 1000-fold than others isolates. The same variability of expression was also observed within commensal isolates from healthy individuals. Apparently, the early selection of six isolates from which the RNA-seq data were obtained included isolates with abnormal expression of this gene, which strongly biased the results described in Chapter 4.

Since we had six nosocomial isolates in our collection that were not found to be associated with infection (either culture contamination or isolates from the skin of hospitalized patients and staff), we performed a secondary analysis, where we included those isolates. As can be seen in Figure 5.5B, the expression levels of *SERP_RS11970* were similar to both infection and community isolates, suggesting that the role of *SERP_RS11970* in the adaptation to blood is transversal to all *S. epidermidis* isolates (used in this study).
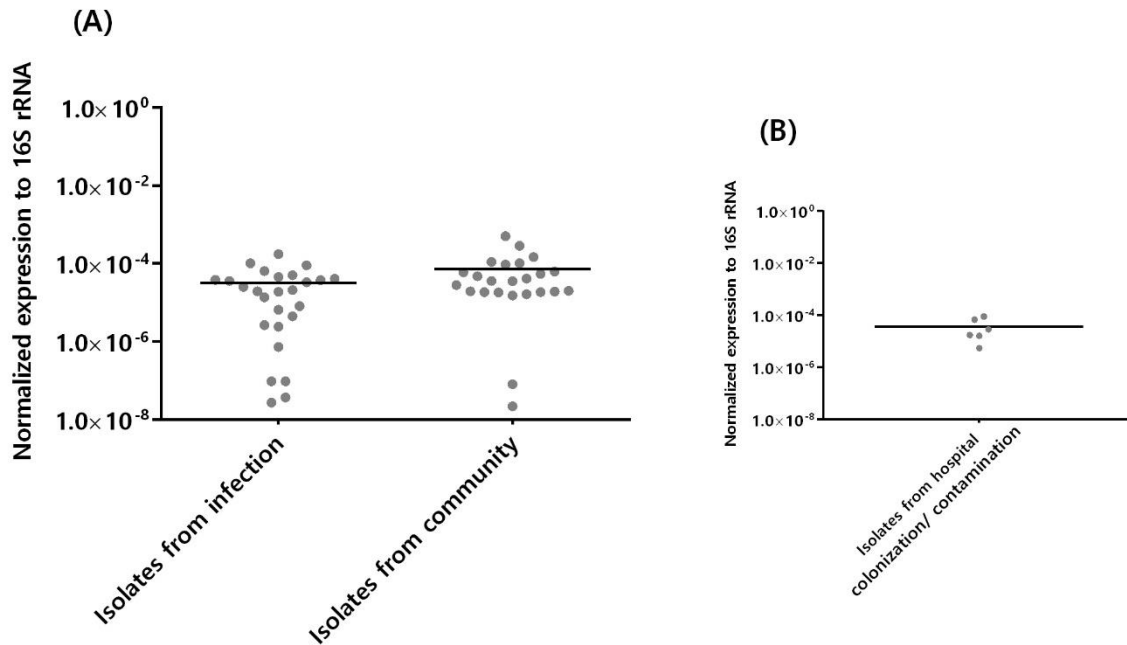
Figure 5.5. Transcription levels of *SERP_RS11970* gene using isolates from infection and from the community (A) isolates from hospital colonization/contamination (B). Each symbol represents the mean of transcription levels of two technical replicates of each isolate. The horizontal line represents the grand mean of the transcription levels of isolates from different origins. Statistical differences between groups were analyzed with one-way ANOVA with Tukey´s multiple comparison test. $p > 0.05$.

We repeated the same analysis by using other target genes highlighted in the previous chapter but, again, the discriminative power of the expression of those genes was lost, as can be seen in Figure 5.6. With these results in mind, our original hypothesis seems not to be validated by the experimental data, at least under the tested conditions. The genes whose expression after co-incubation in blood could discriminate between three distinct clinical isolates, were no longer able to discriminate isolates in a wider collection, including isolates from many European, American, Asian and African countries, although with different representability's from each continent. Epidemiological studies demonstrated that most isolates from the hospital environment seem to be composed by a clonal complex 2, a worldwide disseminated lineage [5] Likewise, *IS256* and *mecA* genes are more prevalent in clinical isolates than in isolates with no contact with the hospital environment [14-16], however, the comparison between gene expression profiles of the genes studied across nosocomial isolates (that included isolates from hospital colonization and contaminants) and commensal isolates did not reveal any difference that could accurately differentiate between infection and contamination. This suggests that *S. epidermidis*, as a species, is able to adapt to human blood in a similar fashion, regardless of being previously associated with an infection or not.
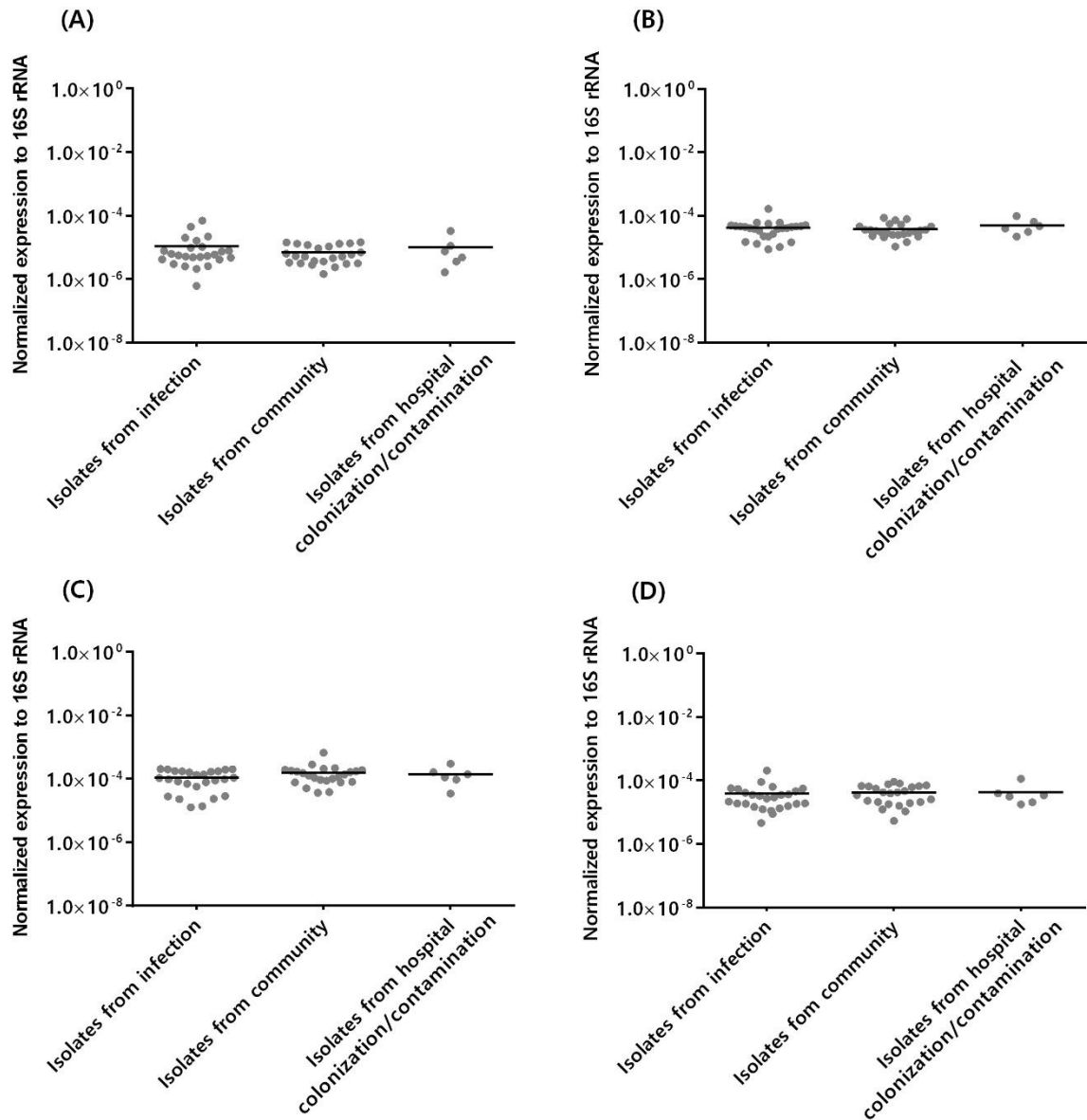
Figure 5.6. Transcription levels of selected target genes using a worldwide collection of *S. epidermidis* isolates. (A) *SERP_ RS10265* **(B)** *SERP_RS05305* **(C)** *SERP_RS06815* and **(D)** *SERP_RS04470* **genes.** Each symbol represents the mean of transcription levels of two technical replicates of each isolate. The horizontal line represents the grand mean of the transcription levels of isolates from different origins. Statistical differences between groups were analyzed with one-way ANOVA and Tukey´s multiple comparison test.

Nevertheless, not being able to demonstrate our original hypothesis, our novel transcriptomic data provided further evidence that, in fact, *S. epidermidis* is an opportunistic pathogen, well adapted to live in the human skin and mucous membranes, and under specific conditions, causing serious nosocomial infections [17]. Furthermore, these findings demonstrated a huge variability in transcriptomic profiles within isolates from the same origin which is consistent with previous research that reported high genetic and phenotypic diversity between different isolates [5, 18, 19]. The inherent

variability of *S. epidermidis* isolates may be related to the need to adapt to different environments, such as hospital and community settings [5].

5.3.4 Bacteria survivability after 4h of incubation

Taken into consideration the transcriptomic data described above, a new hypothesis was proposed: that both commensal and isolates from infection are equally able to survive contact with human blood, and this provides an advantage for the opportunistic nature of this species. The rationale is that to cause infection *S. epidermidis* has to adapt to a new environment, the bloodstream, and develop mechanisms that enable it to overcome the high bactericidal activity of the human blood [20]. Therefore, in order to evaluate the ability of bacteria to survive, the number of surviving bacteria was determined after 4h of incubation in human blood. Contrary to the gene expression studies described in the previous section, for this bacterial survivability assay, human blood was necessary since defibrinated horse blood has its antimicrobial activity compromised because, during the defibrination process, cells with phagocytic capacity are lost [13]. As such, the ability of bacteria to survive was evaluated using 10^5 CFU/mL as initial concentration, in order to closely simulate the number of microbes present in the blood during bacteremia [21].

As shown in Figure 5.7, while there were some isolates that could better survive human blood bactericidal activity, no significant differences were found between the groups tested. The average of all groups was below 10% of survival. Consequently, it appears that all tested *S. epidermidis* isolates have the same ability to adapt to human blood, and eventually cause infection. It should be noted, however, that the bacterial challenged used here was on the high end of what occurs typically in infection [21]. This data further corroborates that no particular trait is exclusively required just for isolates from infection. Nevertheless, this is not the same as saying that there are no particular strains better suited for causing an opportunistic infection. In fact, as mentioned previously, isolates sharing some molecular patterns [14-16] may promote hospital infections, however, they have the same ability to adapt and survive in human blood, and eventually to cause infection, as commensal isolates.
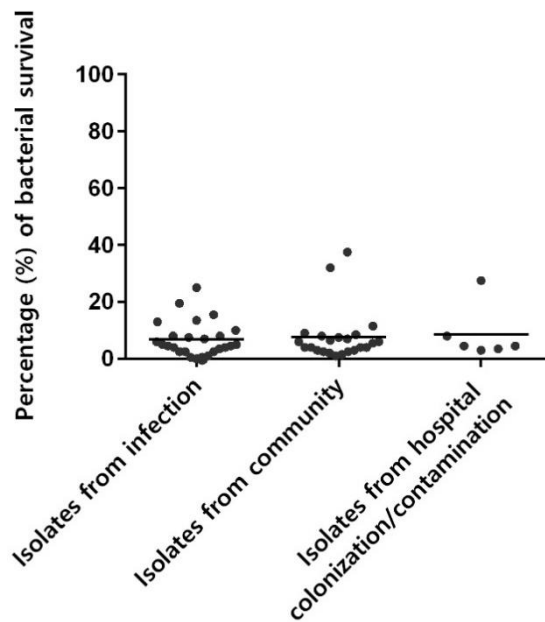
Figure 5.7. The ability to survive in human blood using a worldwide collection of *S. epidermidis* isolates. **Each symbol represents the mean of two to three independent experiments for each isolate.** Statistical differences between groups were analyzed with one-way ANOVA with Turkey´s multiple comparison test. $p > 0.5277$.

5.4 CONCLUSION

The work described in this chapter highlights the importance to validate potentially clinically relevant experimental findings in a wider group of isolates. In Chapter 4, we anticipated that by performing RNA-sequencing with three isolates from each group would provide more robust data for the follow-up studies. Unfortunately, when a larger population of *S. epidermidis* isolates was used, the potential discriminatory of the genes highlighted before was lost. As such, with our experimental design, we were not able to demonstrate that it could be possible to discriminate true infection from blood sample contamination by performing gene expression studies. If the budget was not an issue, RNA-seq could have been done with all 50 isolates (or more), and it is feasible to consider that, under those conditions, different genes could have been highlighted. Nevertheless, our final experiments demonstrated that all isolates had similar survivability in human blood, further strengthening the hypothesis that *S. epidermidis* is an opportunist species [17].

5.5 REFERENCES

1. **Uckay I, Pittet D, Vaudaux P, Sax H, Lew D** *et al.* Foreign body infections due to *Staphylococcus epidermidis*. *Annals of Medicine* 2009; 41(2):109-119.

2. **Christensen GD, Simpson WA, Bisno AL, Beachey EH**. Adherence of slime-producing strains of *Staphylococcus epidermidis* to smooth surfaces. *Infection Immunity* 1982; 37(1):318-326.

3. **Mack D, Siemssen N, Laufs R**. Parallel induction by glucose of adherence and a polysaccharide antigen specific for plastic-adherent *Staphylococcus epidermidis*: evidence for functional relation to intercellular adhesion. *Infection and Immunity* 1992; 60(5):2048-2057.

4. **Cerca N, Martins S, Sillankorva S, Jefferson KK, Pier GB** *et al.* Effects of Growth in the Presence of Subinhibitory Concentrations of Dicloxacillin on *Staphylococcus epidermidis* and *Staphylococcus haemolyticus* Biofilms. *Applied and Environmental Microbiology* 2005; 71:8677-8682.

5. **Miragaia M, Thomas JC, Couto I, Enright MC, de Lencastre H**. Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing data. *Journal Bacteriology* 2007;189 (6):2540-2552.

6. **Freitas AI, Lopes N, Oliveira F, Bras S, Franca A** *et al.* Comparative analysis between biofilm formation and gene expression in *Staphylococcus epidermidis* isolates. *Future Microbiology* 2018; 13:415-427.

7. **Cerca N, Pier GB, Vilanova M, Oliveira R, Azeredo J**. Influence of batch or fed-batch growth on *Staphylococcus epidermidis* biofilm formation. *Letters in Applied Microbiology* 2004; 39:420-424.

8. **Oliveira F, Cerca N**. Antibiotic resistance and biofilm formation ability among coagulase-negative staphylococci in healthy individuals from Portugal. *The Journal of Antibiotics* 2013; 66:739.

9. **Vandecasteele SJ, Peetermans WE, Merckx R, Van Eldere J**. Expression of biofilm-associated genes in *Staphylococcus epidermidis* during in vitro and in vivo foreign body infections. *Journal Infection and Disease* 2003;188 (5):730-737.

10. **França A, Carvalhais V, Maira-Litrán T, Vilanova M, Cerca N** *et al.* Alterations in the *Staphylococcus epidermidis* biofilm transcriptome following interaction with whole human blood. *Pathogens and Disease* 2014; 70(3):444-448.

11. **Franca A, Cerca N**. Plasma is the main regulator of *Staphylococcus epidermidis* biofilms virulence genes transcription in human blood. *Pathogens Disease* 2016; 74(2):27.

12. **Carroll KC, Pfaller MA, Landry ML, McAdam AJ, Patel R** *et al. Manual of Clinical Microbiology, Twelfth Edition*: American Society of Microbiology; 2019.

13. **Wilson JD, Grimes AJ**. Defibrination of normal human blood for in vitro cell studies. *Nature*, Comparative Study 1968; 218(5137):178-180.

14. **Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K** *et al.* The bacterial insertion sequence element IS256 occurs preferentially in nosocomial *Staphylococcus epidermidis* isolates: association with biofilm formation and resistance to aminoglycosides. *Infection Immunity* 2004; 72(2):1210-1215.

15. **Vandecasteele SJ, Peetermans WE, Merckx RR, Rijnders BJA, Van Eldere J**. Reliability of the *ica, aap* and *atlE* genes in the discrimination between invasive, colonizing and contaminant *Staphylococcus epidermidis* isolates in the diagnosis of catheter-related infections. *Clinical Microbiology and Infection* 2003; 9(2):114-119.

16. **Mekni MA, Bouchami O, Achour W, Ben Hassen A**. Strong biofilm production but not adhesion virulence factors can discriminate between invasive and commensal *Staphylococcus epidermidis* strains. *APMIS* 2012; 120(8):605-611.

17. **Otto M**. *Staphylococcus epidermidis*–the 'accidental' pathogen. *Nature Reviews Microbiology* 2009; 7(8):555-567.

18. **Harris LG, Murray S, Pascoe B, Bray J, Meric G** *et al.* Biofilm Morphotypes and Population Structure among *Staphylococcus epidermidis* from Commensal and Clinical Samples. *PLOS ONE* 2016;11 (3):e0151240.

19. **Freitas AI, Lopes N, Oliveira F, Brás S, França Â** *et al.* Comparative analysis between biofilm formation and gene expression in *Staphylococcus epidermidis* isolates. *Future Microbiology* 2018; 13(4):415-427.

20. **Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C** *et al.* Stage-specific gene expression of *Candida albicans* in human blood. *Molecular Microbiology* 2003; 47(6):1523-1543.

21. **Opota O, Croxatto A, Prod'hom G, Greub G**. Blood culture-based diagnosis of bacteraemia: state of the art. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2015; 21(4):313-322.

CHAPTER 6

Conclusion and future work

SUMMARY

In this chapter, the major findings and limitations of this thesis are presented. Furthermore, research questions that should be addressed in the future are also proposed.

6.1 MAJOR OUTCOMES AND THEIR SIGNIFICANCE

This thesis aimed to identify RNA-based molecular markers that could be able to differentiate clinical from commensal isolates and, ultimately, to provide a more accurate diagnostic tool for the discrimination of blood culture contamination from infection. The present thesis started from the hypothesis that *S. epidermidis* isolates from infection or healthy volunteers exhibit different transcriptomic profiles that may allow differentiation between these two groups.

1. Can transcriptome sequencing data provide the identification of molecular diagnostic markers to differentiate *S. epidermidis* true infection from culture contamination?

The analysis and comparison of the transcriptomes of the three clinical and three commensal isolates obtained by RNA-seq, after 2h of incubation in human blood, revealed several potential markers that could allow discriminating clinical from commensal isolates. Unfortunately, when probing the discriminatory potential of the expression of those genes in a wider collection of isolates, we found that transcriptomic profiling was not able to differentiate clinical from commensal isolates. We also included a small group of nosocomial isolates found to be involved in blood culture contamination, and not in true infection, and similarly to the previous observations, we found no differences in the expression of the target genes. As such, at least under our experimental conditions, we could not support the discrimination between isolates contaminating blood culture from infection associated strains, using gene expression analysis.

2. Do clinical and commensal isolates have equal capacity to survive in human blood?

While our initial hypothesis was not demonstrated, our gene expression experiments further confirmed that *S. epidermidis* presents a great versatility that allows isolates from different niches to quickly adapt to a new environment, in this case, human blood, which could potentially lead to an infection. Therefore, our results reinforced the hypothesis that *S. epidermidis* is, in fact, an opportunistic pathogen [1]. In this case, it could be expected that clinical and commensal isolates would be equally able to adapt and survive in human blood. To validate this hypothesis, we compared the ability to survive in human blood of nosocomial isolates from infection, from contaminated cultures or from the community, and no significant differences were found. Therefore, the ability of *S. epidermidis* to cause infection seems to rely on the opportunity to breach the protective layers of the host and its ability to evade and proliferate in the immunocompromised host [1].

86

3. Can we reduce the amount of human blood used in an *ex vivo* model without compromising the analysis of *S. epidermidis* gene expression?

Due to the limitations associated with the availability of human blood donors, as well as due to ethical considerations, we initially addressed the question of to what extreme could we reduce the volume of human blood in each experiment, without impairing the reliability of gene expression studies. Starting with a volume of 1 mL, we were able to further reduce the volume of human blood to as little as 0.18 mL, in a 0.2 mL co-incubation assay, while maintaining the levels of gene expression stable.

6.2 MAJOR LIMITATIONS AND FUTURE WORK

Under our experimental design, we were not able to demonstrate that transcriptomic profiling could be used to differentiate *S. epidermidis* infection-associated isolates from the ones contaminating blood samples. One of the main limitations of this study had to do with the low number of isolates used in the pivotal experiments of RNA-sequencing. This was done taking into consideration the costs of the experiments. Although it was initially considered that three isolates of each condition (3 + 3) would be sufficient to incorporate strain-to-strain variations, as we concluded in Chapter 5, this was not the case, especially due to the large heterogeneity found in *S. epidermidis* isolates.

Another important limitation was the fact that our *ex vivo* human blood model was based on healthy donors. Since *S. epidermidis* causes infection mostly in immunocompromised individuals [2-4], perhaps if using blood from immunocompromised patients, the transcriptomic profile of the isolates could have been altered to the point that more differences in the gene expression profile would have been detected [5].

Another potential limitation is related to the definition of isolates from infection and contamination. This classification was based according to local clinician's evaluation criteria and, thus, the classification of infection or contamination may differ depending on the hospital or geographical area. Hence, there is a small possibility that isolates from infection might have been incorrectly classified as contaminants.

Finally, another limitation was associated with the fact that the collections of isolates used were diverse, with all commensal isolates being from Portugal and all the other isolates from around the world.

While our data further validates the notion that the infections caused by *S. epidermidis* occur in an opportunistic way, future research could address our initial hypothesis but using a different experiment set-up with more focus on the patient than on the bacteria could include:

1. The utilization of an *ex vivo* human blood model using blood from immunocompromised donors;
2. A wider selection of isolates for the RNA-seq analysis studies, and also include more time points;
3. The validation of RNA-seq data should include more pairs of isolates obtained in the same region and in the same timeframe, including isolates associated with infection and blood culture contamination from the same hospital.

6.3 REFERENCES

1. **Otto M**. *Staphylococcus epidermidis*–the 'accidental' pathogen. *Nature Reviews Microbiology* 2009; 7(8):555-567.

2. **Dong Y, Speer CP**. The role of *Staphylococcus epidermidis* in neonatal sepsis: Guarding angel or pathogenic devil? *International Journal of Medical Microbiology* 2014;304(5):513-520.

3. **Nguyen TH, Park MD, Otto M**. Host Response to *Staphylococcus epidermidis* Colonization and Infections. *Frontiers Cell Infection Microbiology*, Review 2017;7:90.

4. **McCann MT, Gilmore BF, Gorman SP**. *Staphylococcus epidermidis* device-related infections: pathogenesis and clinical management. *Journal of Pharmacy and Pharmacology* 2008;60:1551-1571.

5. **Thänert R, Goldmann O, Beineke A, Medina E**. Host-inherent variability influences the transcriptional response of *Staphylococcus aureus* during in vivo infection. *Nature communications*, Article 2017;8:14268.