LEBANESE AMERICAN UNIVERSITY

Comparative Genomic Analysis of *Staphylococcus aureus* CC80 strains.

By

Melhem E. Bilen

A thesis

Submitted in partial fulfillment of the requirements for

the degree of Master of Science in Molecular Biology

School of Arts and Sciences

May 2015

LAU Lebanese American University
الجامعة اللبنانية الأميركية
Lebanese American University
School of Arts and Sciences; Byblos Campus

THESIS APPROVAL FORM

Student Name: Melhem E. Bilen **I.D. #:** 201001355

Thesis Title: Comparative Genomic Analysis of *Staphylococcus aureus* CC80 strains.

Program: M.Sc. in Molecular Biology

Department: Natural Sciences/Biology

School: Arts and Sciences

The undersigned certify that they have examined the final electronic copy of this thesis and approved it in Partial Fulfillment of the requirements for the degree of: M.Sc in Molecular Biology

Thesis Advisor: Dr. Sima Tokajian [redacted]

Thesis Coadvisor: Dr. Georges Khazen [redacted]

Committee Member : Dr. Tarek Nawas

Committee Member : Dr. Roy Khalaf [redacted]

Date: 7/5/2015

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS COPYRIGHT RELEASE FORM

LEBANESE AMERICAN UNIVERSITY NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants to Lebanese American University (LAU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video. You agree that LAU may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that LAU may keep more than one copy of this submission for purposes of security, backup and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant LAU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN LAU, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. LAU will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Name: Melhem E. Bilen

Signature: [redacted]

Date: 7/5/2015

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

PLAGIARISM POLICY COMPLIANCE STATEMENT

I certify that:

- I have read and understood LAU's Plagiarism Policy.
- I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.
- This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: Melhem E. Bilen

Signature: [redacted]

Date: 7/5/2015

ACKNOWLEDMENTS

To begin with, I want to thank the person behind all of what I have become. She is my advisor and friend and always will be. Words will never paint what I have in mind, I just wish that every student would have the same honor as me to make your acquaintance and work with such a wise, dedicated and inspiring person like you. Thank you Dr. Sima Tokajian for being a constant support and believing in me when you had no reason to.

I would also like to thank and show my appreciation to my co-advisor Dr. Georges Khazen. Even though I have only known you for a short period of time, you have proven to be a great listener as well as an encouraging and patient person. I am grateful for all your help.

Not to mention my gratitude to everyone who contributed to my thesis whether academically or experimentally. Thank you Tamara Salloum, Dr. Roy Khalaf, Dr. Tarek Nawas, Maya Farah and Helena Bou Farah, your help is much appreciated.

Last but not least, I am blessed to have a supporting and motivating family whether blood related or not. I'm privileged to have you in my life. Thank you mom, dad, my uncle, Christel Dagher and Joe Ibrahim for always being there no matter the circumstances.

I dedicate my work to my beloved deceased aunt Wadad M. Bilen

Comparative Genomic Analysis of *Staphylococcus aureus* CC80 strains.

Melhem E. Bilen

ABSTRACT

Staphylococcus aureus is a highly adaptive and versatile gram-positive bacterium and is considered one of the most isolated human pathogens and the most common cause of skin and soft tissue infections (SSTIs). The ongoing evolution of antimicrobial resistance namely the particularly methicillin-resistant strains of *S. aureus* (MRSA) has complicated the treatment of infections with such isolates. Global transmission of MRSA has been linked to international travel. However, in countries with a high prevalence of MRSA, imported cases are not easily distinguished from the domestic background prevalence. In this study we have selected ST80 MRSA strains isolated from clinical samples from Lebanon and Jordan for genome sequencing. DNA extracted from all isolates was prepared for sequencing, pooled together and run on Illumina MiSeq for paired-end 250-bp sequencing. Those isolates, found to be ST80, carried several virulence factors, namely PVL that causes a wide variety of diseases such as skin and mucosal lesions including necrotic hemorrhagic pneumonia. Additionally, intracellular adhesion proteins and clumping factors involved in biofilm formation and adhesion respectively were identified. Fibronectin binding proteins that play a role in invasion were expressed in all of our isolates. Moreover, exfoliative toxin, eta, which is the causative agent of staphylococcal scalded skin syndrome, was also detected. Nevertheless, different plasmids were identified in those isolates carrying several resistance genes towards Tetracyclines and Beta-lactams. Also, some isolates were shown to be resistant towards Lincosamide, Straptogramin and Macrolides. Finally, integrated phages, such as phi7401PVL and JS01 were identified in most of the isolates along with several CRISPR sequences that can be used in targeted therapies. Knowledge of epidemic MRSA clones can help in the development of effective strategies to aid in

optimizing treatment, and revealing the mode of pathogenicity. The detection of rare SCC*mec* cassettes along with rare combination of toxin encoding genes, resistance factors and other virulence determinants could be indicative of MRSA clones being imported and warrants further studies to better understand and deal with such strains.

Keywords: *Staphylococcus aureus,* MRSA, ST80, Virulence determinants, Resistance genes.

Table of Contents

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS

ACME arginine catabolic mobile element

AUBMC American University of Beirut Medical Center

CA-MRSA community acquired MRSA

Cas CRISPR-associated

CHIP Chemotaxis Inhibitory Proteins

CRISPR Clustered Regularly interspaced short palindromic repeats

EAP Extracellular Adherence Proteins

EMP extracellular matrix proteins (EMP)

ffnb fibronectin binding protein

fib fibrinogen binding proteins

Fn Fibronectin

hVISA heterogenous vancomycin-intermediate *S. aureus*

ica intracellular adhesion proteins

MRSA methicillin-resistant strains of S. *aureus*

MSCRAMM microbial surface components recognizing adhesive matrix molecules

NGS New generation sequencing

PBP-2a penicillin-binding protein 2a

PMS phenol soluble modulin

PVL Panton-Valentine leukocidin

REP replication protein

SCC staphylococcal chromosomal cassettes

SNP single nucleotide polymorphism

SSTI skin and soft tissue infections

STSS staphylococcal toxic shock syndrome

TSA Trypticase Soy Agar

TSST-1 toxic shock syndrome toxin-1

VISA vancomycin-intermediate *S. aureus*

VRSA vancomycin-resistant *S. aureus*

WGS Whole genome sequencing

STSS staphylococcal toxic shock syndrome

Chapter One

INTRODUCTION

S. aureus is a common gram-positive bacterium known for its spherical shape and clusters. This bacterium resides in the epithelial lining of the upper respiratory tract and the skin (Abdelnour et al., 1993). It was discovered by a Scottish doctor Alexander Ogston during a knee operation, and studies showed that more than 20% of the human population are permanent carriers of that bacterium (Ogston, 1889). Fortunately, these diseases are mostly treatable and could be therefore prevented (Chambers et al., 2009; Moran et al., 2006). The ability of *S. aureus* to cause diseases is multifactorial, a combination of many virulent factors including toxins, secreted exoproteins, and cell surface-associated adhesins (Young et al., 2012). *S. aureus* resistance against a wide variety of antimicrobials is another problem and is increasing worldwide. Thus, the need to decipher the DNA is critical in order to follow the evolution of mutations, resistant genes and virulence factors. Scientists obtained the capability to decode the genomic information of any biological system with the emerging of capillary electrophoresis based on Sanger sequencing. However, this approach has still its limitations in for instance information processing, measurability, rapidity and resolution, which always prohibit scientists from collecting the fundamental needed data (Fricke et al., 2013). Next-Generation Sequencing (NGS) on the other hand, enables the comprehensive examination and interpretation of genomes at low costs, higher speed and increased throughputs. It additionally, permits easier clinical diagnostics, effective therapy and prenatal diagnosis (Koboldt et al., 2013). Such recent technologies are rapidly developing and require new protocols for preparing DNA libraries and establishing updated and easy approaches to analyze data and revise the experimental design (Köser et al., 2012).

In our study, several *S. aureus* strains were obtained from the American University of Beirut Medical center (AUBMC) and from Jordan University hospital with the aim of their molecular characterization. Genome sequencing was performed to

identify the genes responsible for resistance, virulence, and pathogenicity as well as to establish a comparative analysis and phylogeny.

Objectives:

- Screening the bacterial genome for the presence of virulence genes and pathogenicity islands.
- Identifying genes responsible for resistance and understanding resistance mechanisms.
- Perform a comparative analysis between the different *S. aureus* strains and their respective worldwide sequenced reference strains.
- Trace the evolutionary lineage among the different isolates through the construction of a single nucleotide polymorphism (SNP) based phylogenetic tree.

Chapter Two

LITERATURE REVIEW

Member of the Firmicutes phylum, *Staphylococcus aureus* is frequently found on the skin and the epithelial surface of the respiratory tract. Known to be pathogenic, this bacterium is the origin of many diseases including boils (cutaneous infection), sinusitis (pulmonary infection), urinary tract diseases and several other illnesses. Disease-causing strains transmit the infection through the production of toxins or the expression of surface proteins. The evolution of *S.aureus* genome is in a race with scientists who are constantly trying to discover new antibiotics, track their origin and limit their evolution. Thus, typing and whole genome sequencing of the *S. aureus'* genome could be a critical step towards the discovery of all the genome contents and the determination of new therapeutic methods.

2.1. Virulence factors:

S. aureus is known for causing a variety of infections and it does so with the aid of virulence factors. Virulence factors enable the bacterium to avoid the immune system, invade the host, release toxins that harm the host and adhere to surfaces. The majority of virulence factors can be grouped under two categories:

2.1.1. Cell surface components:

S. aureus has a group of surface proteins known as microbial surface components recognizing adhesive matrix molecules (MSCRAMM) that enable its attachment to the host and instigate colonization thus causing an infection. Virulence factors associated with the cell wall of *S. aureus* include capsular polysaccharides, which are found in around 90% of clinical isolates. Staphylococci capsules play a major role in virulence by blocking neutrophils from phagocitizing them (Johannessen et al., 2012**).** They also improve the colonization of the bacteria especially when it comes to mucosal

surfaces. In addition to capsules, *S. aureus* has staphyloxanthin, which is a pigment that also helps it in evading neutrophils. If the biosynthesis of staphyloxanthin were to be inhibited, then this would render *S. aureus* vulnerable to innate immune responses (Clauditz et al., 2006). Furthermore, *S. aureus* has a wide variety of surface proteins such as fibronectin binding proteins (Fnb), collagen adhesion, protein A and clumping factors among others. These proteins enable the bacterium to adhere to the proteins of the host and establish an infection as well as hide the microorganism from the host's immune system (Eidhin et al., 1998); Staphylococcal protein A (*spa*) attaches to the Fc portion of the antibody to interfere with opsonization (DeDent et al., 2007).

2.1.2. Secreted components:

Virulence factors secreted by *S. aureus* constitute four major categories: toxins forming pores, exoenzymes, superantigens and various proteins (Schlievert et al., 2010). Leukocidin is a pore-forming toxin, at low concentration, it causes apoptosis while at high concentrations it causes cell lysis (Genestier et al., 2005). Lipase on the other hand is an exoenzyme found in *S. aureus* and its role is to degrade fatty acids (Dinges et al., 2000). Furthermore, superantigens are exotoxins produced by *S. aureus* and have the ability to engender several human diseases. Twenty different superantigens have been extracted from *S. aureus* and over 60 percent of *S. aureus* isolates have a minimum of one superantigen (Klutymans et al., 1997). In comparison to antigens that activate around 0.001% of T-cells, superantigens are capable of activating 5-30% of T-cells. The activation of T-cells instigates the release of chemokines and cytokines from both T-cells and antigen presenting cells such as dendritic cells and macrophages. Those chemicals then will be responsible for establishing the symptoms of staphylococcal diseases such as staphylococcal toxic shock syndrome (STSS) (Ortega et al., 2010; Bachert et al., 2002).

2.2. Pathogenesis:

Infections are more prominent to occur when there is an open wound. A viral infection in the upper respiratory pathway would impair the mucosal linings and therefore render the host susceptible to *S. aureus* pneumonia (Rubinstein, 2008). Once *S. aureus* is exposed to the host tissue past the skin barrier and mucosal surface, its virulence factors will be up regulated. On the other hand, phagocytes usually reside in the epithelial and mucosal tissue, and trigger the immune system in case of tissue injury or bacterial invasion. Similarly, the host has pattern recognition molecules that sense the lipoproteins and the peptidoglycan layer of *S. aureus,* which facilitates inflammatory signaling including the recruitment of macrophages and neutrophils (Kim et al., 2011).

S. aureus is also known to thrive well both outside and inside the cells of the host. *S. aureus* must avoid antibodies and complement component proteins in order for it not to be opsonized. To do so, *S. aureus* expresses on its wall clumping factor A also known as protein A, as well as other complement inhibitors; all of which can hinder opsonization (Gordon et al., 2008). In addition, *S. aureus* can hide inside epithelial cells and macrophages, but not in neutrophils since they are more challenging as evidenced by *S. aureus* infections being more invasive in patients with dysfunctional neutrophils. This however, doesn't imply that *S. aureus* do not have strategies to resist killings by neutrophils. One of its strategies is to secrete Extracellular Adherence Proteins (EAP) and Chemotaxis Inhibitory Proteins (CHIP) both of which interfere with the ability of neutrophils to attach to epithelial cells through interacting with intercellular adhesion moleculaes (ICAM-1) and impede neutrophils from recognizing chemotactic factors, respectively (Foster, 2005).

Nevertheless, once the neutrophils get to the site of infection, they release antimicrobial products such as reactive nitrogen species, lysozyme, and proteases. *S. aureus* protects itself from reactive oxygen species by releasing antioxidant enzymes such as catalase and superoxide dismutase (Liu, 2009). *S. aureus* on the other hand, resists antimicrobial peptides that are aimed at negatively charged bacteria by altering its own surface charge. This pathogen also neutralizes and degrades antimicrobial peptides with staphylokinase and aureolysin, respectively (Sieprawska-Lupa et al., 2004).

Another strategy used by *S. aureus* to evade neutrophils is through releasing toxins consisting of two components that are specific to human and not mouse cells. Till date,

the function of those toxins were not clearly elucidated. A combination of bacterial peptides known as phenol soluble modulin (PSM) have been also identified, and their role in virulence was mainly associated with CA-MRSA strains covering skin infections (Graves et al., 2010).

Other than evading the immune system of the host, it is essential for the bacterium to obtain its nutrients especially iron in order to survive inside the human host. When an infection occurs, iron in the serum is made difficult to attain by binding to host proteins, beside the fact that 95% of iron is secluded inside host cells. When iron however is scarce, *S. aureus* secretes substances that have high affinity to iron like staphyloferrin and aureochelin. Similarly, when iron concentrations are low, *S. aureus* instigate the transcription of iron retrieval mechanism, which enables the bacterium to capture the haptoglobin and heme on the surface of the cell, then carry the complex into the cytoplasm and finally degrade the heme oxidatively (Hammer et al., 2011).

There are also other virulence factors worth mentioning including the ability to form biofilms. Biofilms allow *S. aureus* to stick to plastic, counteract the defense of the host or antibiotics and produce variants of small colonies that allow it to survive when there are severe conditions by entering a metabolically inactive state (Liu, 2009).

2.2.1. MRSA Pathogenesis

It is evident that MRSA has distinct epidemiology especially when it comes to mortality and morbidity. It is separated into HA-MRSA and CA-MRSA, both of which are considered to be genotypically different targeting distinct communities but yet overlapping and causing diverse diseases (Evans, 2008).

2.2.2. HA-MRSA:

HA-MRSA has staphylococcal chromosomal cassettes (SCC*mec*) type I as well as SCC*mec* type II and III, which encode for a single or even multiple antibiotic resistance (Moroney et al., 2007), and hence capable of surviving in environments where antibiotics are frequently being used (Enright et al., 2012).

It's noteworthy that once HA-MRSA is outside the healthcare environment, it will rarely cause diseases in humans unless they are predisposed. As a consequence, it is indicated that HA-MRSA is one of *S. aureus* less vigorous strains and could only thrive where the competition is dependent on antibiotic pressure. Furthermore, studies performed showed that HA-MRSA are more susceptible to neutrophil killings and are less pathogenic when injected into mice systematically (Voyich et al., 2005). When comparing HA-MRSA and CA-MRSA strains, it was proven that HA-MRSA expressed reduced levels of PSM peptides, thereby signifying a probable defect in the regulation of virulence in HA-MRSA (Wang et al., 2005). The expression of most virulence factors is controlled by the global regulator (*agr*), which is thought to be a prime pathogenesis factor in *S. aureus*. Clinical HA-MRSA isolates are arg^- or arg^+ genotypes. These genotypes (arg^- or arg^+) explain the non-pathogenicity of HA-MRSA in immunocompetent individuals. However, the same genotype might be favorable for the survival of HA-MRSA in healthcare environment (Shopsin et al., 2008); arg^- enhances biofilm development and proliferation especially on plastic tubes (Vuon et al., 2000).

2.2.3. CA-MRSA:

CA-MRSA strains were alleged to be more virulent than HA-MRSA since they cause unconventionally severe infections like myositis and necrotizing pneumonia. However, due to insufficient comparative studies on CA-MRSA and methicillin sensitive *S. aureus*, it cannot be concluded that all CA-MRSA clones are more virulent (Bardley, 2005). Yet, there is one CA-MRSA clone that has been proven to be exceptionally successful, known as USA300. USA300 has been found in Europe, Canada and the United States and reports show that it is related to severe infections in soft tissues, bone and skin (Miller et al., 2008).

The virulence factor of CA-MRSA being most studied is the PVL. PVL was detected in the phage type 80/81 epidemic *S. aureus* strain. PVL, a two-component toxin, when injected into mice or rabbits, evident signs of apoptosis, neutrophil cytolysis, inflammation and necrosis were visible. Another important virulence factor is the type I arginine catabolic mobile element (ACME). It is speculated that ACME is horizontally

transferred from omnipresent skin commensal *S. epidermidis*. The two most significant gene clusters found in ACME are *opp-3* and *arc*. *opp-3* is an ABC-transporter, which functions in improving the survival of the bacteria on the skin through: resisting antimicrobial peptides, improving eukaryotic cell adhesion, and enhancing the uptake of peptide nutrients. Moreover, *arc* is an arginine deaminase system, which breaks down L-arginine in order to provide the bacterium with ATP and raise the acidic pH of the human skin to one that is more convenient for bacterial colonization. Therefore, once *S. aureus* acquires ACME, then it might enhance the ability of CA-MRSA to colonize and consequently induce a skin infection once the skin barrier is disrupted. Evidence suggests that CA-MRSA is better at colonizing in comparison to HA-MRSA and MSSA; however, no direct evidence indicates that ACME actually aids in skin colonization (Liu, 2009).

2.3. Resistance evolution and mechanism:

S. aureus was able to develop resistance mechanisms toward most of the antibiotics developed since 1950 after introducing Penicillin till 2002 when the first Vancomycin resistant *S. aureus* was reported (Figure 1).

2.3.1. Penicillin resistance

2.3.1.1. History

Before the usage of the first *S. aureus* antibiotic, the mortality rate caused by these bacteria reached 80% (Skinner, 1941). In 1940, penicillin was introduced by Alexander Fleming as the first antibiotic against *S. aureus* and showed high efficiency in treating patients. Unfortunately, the first penicillin resistant strain was discovered in the hospitals in 1942. It spread quickly in the community and was responsible for 80% of the cases in 1960 (Deurenber et al., 2008).

2.3.1.2. Resistance mechanism:

Studies continued in order to discover the resistance mechanism of the *S. aureus* toward penicillin. *S. aureus* was able to inactivate penicillin through an enzyme called β-lactamase which is encoded by *blaZ* gene and regulated by *blaR1* and *blaI*. The pathway leading to β-lactamase synthesis is activated after the introduction to β-lactams through the cleavage of the *blaR1* and *blaI* proteins (Lowy, 2003).

2.3.2. Methicillin resistance

2.3.2.1. History

Methicillin was released in 1959 and used to treat patients carrying penicillin resistant *S. aureus* strains. However, methicillin resistant strains evolved in a short period of time and the first case was reported in a British hospital. MRSA spread quickly among different populations with severe symptoms and high number of nosocomial infections (Jevons, 1961).

2.3.2.2 Resistance mechanism:

mecA gene codes for the factors responsible for the resistance against methicillin in MRSA. It's a part of the staphylococcal cassette chromosome *mec* that can be divided into five types: I, II, III, IV, and V. Type I, II and III were found to be associated with HA-MRSA while type IV and V were associated with CA-MRSA strains. Recently, novel SCC*mec* types (VI to XI) have been reported along with many subtypes (Turlej et al., 2011). *mecA* gene conferred resistance through the production of a modified plasma membrane bound enzyme called penicillin-binding protein 2a (PBP-2a). PBP-2a was able to perform its job having low binding affinity towards β-lactam antibiotics (Hartman et al., 1984).

2.3.3. Quinolone-Resistant *S. aureus*

2.3.3.1. History:

Fluoroqinolones were first introduced in 1980s to treat gram-positive bacteria with *S. aureus* quickly developing resistance against this group of antibiotics (kirst et al., 1998).

2.3.3.2. Resistance mechanism:

The mechanisms of resistance toward fluoroquinolones are based on mutations in the genes coding for topoisomerase IV or DNA gyrase, or on the presence of multidrug efflux pump. DNA gyrase is responsible for relieving DNA supercoiling and topoisomerase IV disentangles accumulated DNA. However, any mutation involving a single or multiple amino acids especially in the region of the enzyme-DNA complex results in affinity reduction. Mainly, mutations in the ParC subunit of topoisomerase IV and GyrA subunit in gyrase are responsible for this resistance. *NorA* is an example of an efflux pump that was first discovered in 1986 in a fluoroquinolone-resistant strain. *NorA* is able to secrete molecules similar to fluoroquinolones, like norfloxacin and ciprofloxacin. Therefore, the over expression of the *NorA* gene through a mutation in the promoter region, or a mutation in *NorA* gene regulators, or a mutation in any inducible or constitutive region, will lead to an increase in resistance. Moreover, the accumulation of resistant bacteria along with the low quinolone concentrations applied lead to an environment that enhances the growth of resistant mutants (Jacoby et al., 2005).

2.3.4. Vancomycin resistance:

2.3.4.1. History

Vancomycin-resistant strains can be divided into 3 main groups: vancomycin-intermediate *S. aureus* (VISA), heterogenous vancomycin-intermediate *S. aureus* (hVISA), and high-level vancomycin-resistant *S. aureus* (VRSA) (Jacoby et al., 2005). In 1997, VISA was first discovered in Japan and then in other countries such as France, Brazil, and the United States (Smith et al., 1999). These strains showed resistance toward all glycopeptide antibiotics and had a phenotype characterized by a thick cell wall believed to reduce the efficiency of vancomycin penetration into the cell.

In 2002, VRSA strain was first identified from a catheter of a diabetic patient in Michigan carrying the *mecA* and *VanA* genes. *VanA* gene was shown to be encoded on a transposon Tn1546 in VRSA isolates, which conferred vancomycin resistance in enterococci (Chang et al., 2003; Weigel et al., 2003). Until 2010, eight patients suffering from VRSA were reported in the US and two others from Iran and India (Gould, 2010).

2.3.4.2. Mechanism of resistance:

Theories about the resistance of *S. aureus* to vancomycin started pointing towards the formation of a modified cell wall. Later on, it was proven that *S. aureus* showed an alteration in peptidoglycan biosynthesis resulting in a thicker cell wall with an irregular shape along with a decrease in cross linkage between peptidoglycans (Hanaki et al., 1998). The cell wall contained more D-Ala-D-Ala residues that bind and trap vacomycin preventing it from reaching its target (Hanaki et al., 1998). Vacomycin resistance is also due to the conjugal transfer of *vanA* gene from the vancomycin-resistant *Enterococcus faecalis* and other enterococci (Showsh et al., 2001).

2.4. Whole genome sequencing (WGS):

2.4.1. Overview:

New Generation Sequencing (NGS) enables the comprehensive examination and interpretation of genomes, trascriptomes and interactomes at lower costs, higher speed and increased throughputs (Mardis, 2009). Also, sequencing of DNA permits easier clinical diagnostics including: effective therapy, identification, and prenatal diagnosis (Gregg et al., 2014).

2.4.2. WGS of *S. aureus*:

Based on WGS, *S. aureus* was found to carry genomic islands with polymorphisms among different strains and with several genes encoding virulence factors (Baba, 2013).

Moreover, WGS was also used to study the antimicrobial resistance properties of 501 unrelated *S. aureus* and compare it to results obtained from routine tests. The study showed that WGS was way more sensitive and accurate than any other commercial test (Stegger et al., 2014). Other findings targeted the pathogenicity islands using WGS in *S. aureus* strains, and three new pathogenicity islands were discovered including a toxic-shock-syndrome toxin island family, exotoxin islands, and enterotoxin islands (Kuroda et al., 2001). In the UK, studies on MRSA outbreaks in a childcare center were done using WGS. The data showed that MRSA can be carried by a staff member and permitted the outbreak to endure for phases without any infection announcement (Harris et al., 2013).

Additionally, 3A divergent *mecA* homologue (*mecA* (LGA251)) was discovered in the LGA251 genome located in a novel SCC*mec* element designated type-XI (Alvarez et al., 2011).

This technique also helped to study polymorphisms in different strains such as RN4220 and NCTC8325 (Nair et al., 2001), or to compare livestock associated methicillin resistant strains isolated from Europe and Canada (Golding et al., 2010).

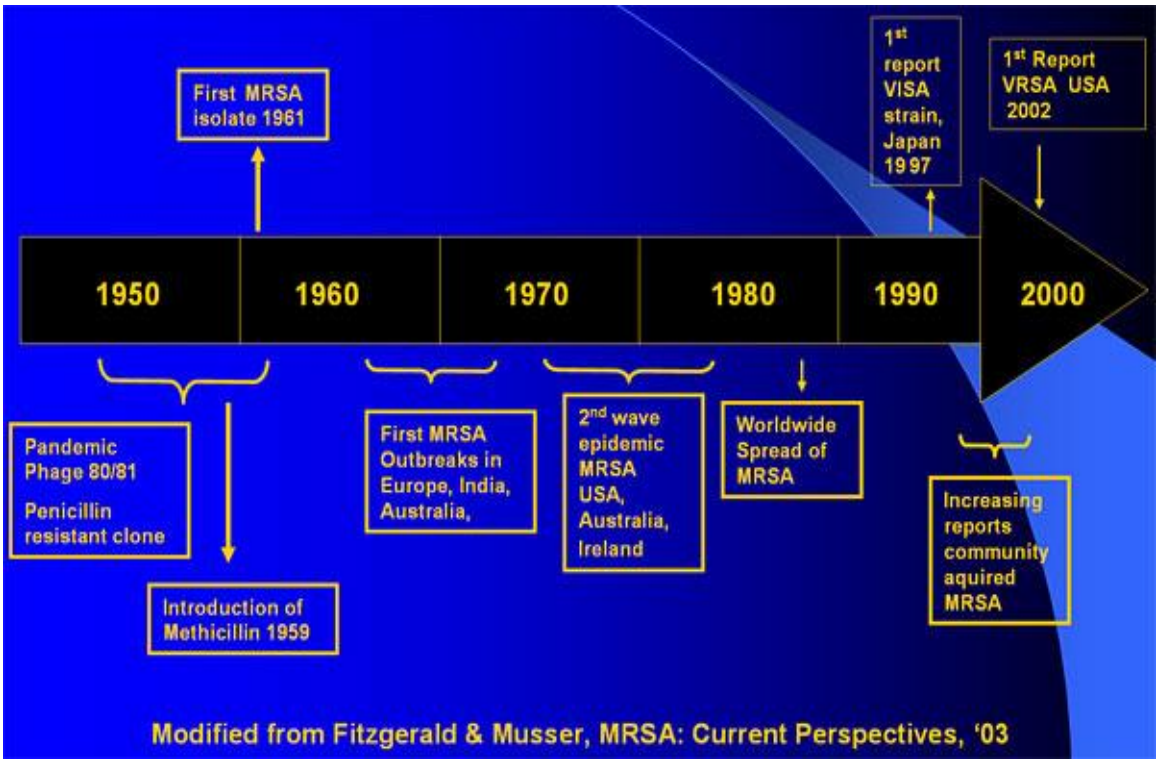

Figure 1: Evolution of antibiotics resistance of *S. aureus* from 1950 to 2000.

2.5. Infection and Diseases:

S. aureus can infect the majority of organ systems, studies showed that more than 20% of the human population is a permanent carrier of that bacteria (Young et al., 2012). This bacterium is most often spread to others by contaminated hands and biomedical materials. The skin and mucous membranes are usually an effective barrier against infection. However, if these barriers are breached (e.g., skin damage due to trauma or mucosal damage due to viral infection) *S. aureus* may gain access to underlying tissues or the bloodstream and cause infection (Odell, 2010). The spread of the infection ranges from being limited such as furuncle or carbuncle to expanding to other areas such as impetigo and cellulitis (Dryden, 2010). Furthermore, when the bacteria reaches the bloodstream, it causes bacteremia and hence it disperses to different sites of the body where it causes infections such as osteomyelitis, visceral endocarditis and pneumonia among others that might cause severe damage organ or kill the host. It should be noted that *S. aureus* strains also synthesize exotoxins and enzymes that most probably cause or enhance the asperity of certain diseases.

Chapter Three

MATERIALS AND METHODS

3.1. Ethical Approval

Ethical approval was not required as clinical isolates were collected and stored as part of routine clinical care. Clinical isolates and patient records/information were anonymized and de-identified prior to the analysis.

3.2. Bacterial isolates

This study was conducted on *S. aureus* bacterial isolates, collected from the AUBMC (MB6, MB7, MB9 and MB12) and Jordan University Hospital (MB42). These samples were isolated from wound and were cultured overnight on Trypticase Soy Agar (Bio-Rad, USA) medium for subsequent experimental work.

3.3. DNA extraction

Bacterial DNA was extracted using the Nucleospin Tissue kit (Macherey-Nagel, Germany) and following the manufacturer's instructions.

3.4. Genome sequencing:

After fluorescent DNA quantification using Qubit 3.0 fluorometer (Life technologies, Carlsbad, CA), the samples are prepared for multiplexed, paired-end sequencing on a MiSeq Desktop Sequencer (Illumina, Inc., San Diego, CA). For each isolate, 1 ng of input DNA was used for indexing, and amplification steps were done using by the Nextera XT library prep kit (Illumina, Inc., San Diego, CA) and following the

manufacturer's instructions. The subsequent clean up steps were performed using the AMPure XP PCR purification beads (Agencourt, Brea, CA) and again following instructions supplied by the Illumina manual. The resulting individual DNA libraries with fragments ranging from 500-1000bp were quantified by quantitative PCR on a CFX96 (Bio-Rad, USA) in triplicate at two concentrations, 1:1,000 and 1:2,000, using the Kapa library quantification kit (Kapa Biosystems, Woburn, MA). Based on the individual library concentrations, equimolar pools of no more than 12 indexed libraries were prepared at a concentration of at least 1 nM using 10 mM Tris-HCl (pH 8.0) and 0.05% Tween 20. To ensure accurate loading onto the flow cell, the same quantification method was used to quantify the final pools. Finally the pooled paired-end libraries were sequenced on a MiSeq Desktop Sequencer to a read length of at least 250 bp.

3.5. Analysis of sequencing results

3.5.1. Genome assembly

Genome assembly was done *de novo* using A5 with default parameters. This pipeline automates the processes of data cleaning, error correction, contig assembly, scaffolding, and quality control (Tritt et al. 2012). Reference mapping to the closest reference strains and *de novo* assembly were performed using the MIRA assembler v4.0.2 (Chevreux et al., 1999).

3.5.2. Genome annotation using RAST

The RAST server (http://rast.nmpdr.org) was used to annotate the obtained *de novo* assemblies, identification of protein encoding, rRNA, and tRNA genes, as well as assigning gene functions and predicting subsystems present in the genome (Aziz et al., 2008; Overbeek et al., 2014).

3.5.3. Determination of resistance genes, virulence genes, and pathogenicity

Resistance and virulence genes found in the bacterial genomes were determined by mapping the WGS data to online databases. The SEED viewer service from RAST proved to be very useful in extracting all the genetic elements found on the sequenced bacterial genomes. It was used, combined with the web tools provided by the Center for Genomic Epidemiology (CGE) website (www.genomicepidemiology.org) to get a detailed list of the genetic elements in question. The ResFinder 2.1 web server was used to identify acquired antimicrobial resistance genes present on the bacterial genome, the job was run with the website's default parameters (Zankari et al., 2012). ResFinder detects the presence of resistance genes, but does not confirm their functional integrity and expression. Resistance to acquired mutations in housekeeping genes, is another feature, the service misses. Hinging on the ResFinder hits, a resistance profile was established using phenotypes from previous published studies. Virulence genes were determined using the VirulenceFinder 1.2 service from the same website (Joensen et al., 2014).

3.5.4. Multi-locus sequence typing (MLST) and plasmid detection

MLST typing of the isolates was done by mapping the WGS data to an online database using CGE's MLST 1.7 server (Larsen et al., 2012). Plasmid detection was also achieved using the website's PlasmidFinder 1.2 web service with a 95% selection threshold (Carattoli et al., 2014).

3.5.5. Phage detection

Phage detection was done using the publically available Phage Search Tool (PHAST) (http://phast.wishartlab.com/index.html) (Zhou et al., 2011). This tool provided us with site of phage integration and assigns a phage family. Sequences were annotated using RAST.

3.5.6. Circular visualization

Circular genome visualization was done using two different tools. Comparative figure were generated using the CGView server at the publically available Stothard Research Group website (http://stothard.afns.ualberta.ca/cgview_server/) (Grant & Stothard, 2008). Informative figures for each sample were also generated using DNAPlotter release 1.11 (Carver et al., 2009).

3.5.7 SNPs detection and construction of a phylogenetic tree

SNPs calling, filtering and alignment were done using CSI Phlogeny 1.0a web service (https://cge.cbs.dtu.dk/services/CSIPhylogeny) (Kaas et al., 2014). Maximum likelihood Tree visualization was done using FigTree v1.4.2.

3.5.8. Clustered regularly interspaced short palindromic repeats (CRISPR) detection

CRISPRFinder was used for detection of CRISPRs in locally created data and consultation of CRISPRs present in the database. Also, information on the presence of CRISPR-associated (*cas*) genes were collected (Grissa et al. 2007).

Chapter Four

RESULTS

4.1. Genomes overview:

General features of the isolates were determined using the RAST server and were tabulated in a comparative way (Table 1).

Table 1: General features of the isolates in terms of size, number of contigs, GC content, contigs size, number of coding sequences, number of subsystems and RNAs.

	MB46	MB6	MB7	MB9	MB12
Genome Size (bp)	2,825,714	2,837,575	2,835,620	2,860,243	2,819,327
No of Contigs	32	86	48	32	38
GC mol %	32.70%	32.9	32.8	32.8	32.7
Shortest contig size	511	498	499	532	517
Longest contig size	432197	467917	381927	440641	395327
Number of coding sequence	2624	2619	2651	2667	2597
Number of subsytems	403	401	401	402	403
Number of RNAs	83	83	77	80	79

4.2. MLST typing:

All isolates taken in this study were identified using MLST 1.7 and showed to be ST80 (Table 2).

Table 2: The results obtained from MLST 1.7 server using CGE were tabulated in a comparative way in order to check similarities between the different strains. All five strains were identified as ST80 with a 100% identity for seven loci *arc, aroe, glpf, gmk, pta, tpi, yql* (Table 2).

Samples	*Locus*	% Identity	HSP Length	Allele Length	Gaps	Results
MB46	*arcc*	100	456	456	0	
MB6	*aroe*	100	456	456	0	
MB7	*glpf*	100	465	465	0	
MB9	*gmk*	100	417	417	0	
MB12	*pta*	100	474	474	0	
	tpi	100	402	402	0	
	yqil	100	516	516	0	**ST-80**

4.3. Virulence factors

Virulence factors obtained using VirulenceFinder 1.2 service, were grouped in a table showing the genotypic differences between the isolates and between CC80 reference strain (ST_11819-97). All isolates showed to have the same set of genes that code for intracellular adhesion proteins (*IcaA*, *IcaB*, *IcaC*, *IcapR*). Additionally, the Panton-Valentine leukocidin (*LukF-PV, lukS-PV*), exotoxins, complement and fibrinogen binding proteins (fib) were detected in all the isolates (Table 3).

Table 3: This table shows a summary of the detected virulence factors, a more comprehensive table about virulence factors can be found in Annex I.

Isolates	**MB46**	**MB6**	**MB7**	**MB9**	**MB12**	**CC80 Reference Strain**
Protein function	**Virulence factors**	**Virulence factors**	**Virulence factors**	**Virulence factors**	**Virulence factors**	**Virulence factors**
Intracellular adhesion proteins	*IcaA,IcaB, IcaC IcaD,Icap R*	*IcaA,IcaB, IcaC IcaD,Icap R*	*IcaA,IcaB, IcaC IcaD,Icap R*	*IcaA,Ica B,IcaC IcaD,Ica pR*	*IcaA,Ica B,IcaC IcaD,Ica pR*	*IcaA,IcaB, IcaC IcaD,Icap R*
Exfoliative toxin A	*eta*	*eta*	*eta*	*eta*	*eta*	*eta*

lukF-PV, lukS-PV	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*
Clumping factor ClfB, fibrinogen binding protein	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*
Fibronectin binding protein	*FnbA and FnbB*	*FnbA and FnbB*	*FnbA and FnbB*	*FnbA and FnbB*	*FnbA and FnbB*	*FnbA and FnbB*

4.4. Resistance genes:

Resistance genes in all isolates and CC80 reference strain ST_11819-97 were identified using ResFinder. Different resistance patterns were detected, with all being tetracycline and β-lactam resistant. MB46 was only resistant to lincosamide, streptogramin B and macrolides resistant, whilw both MB9 and CC80 reference strain were also resistant to the macrolides. All isolates except MB7 were aminoglycoside resistant, and four isolates (MB46, MB6, MB9 and MB12) were resistant to fusidic acid (Table 4).

Table 4: Resistance patterns in ST80 isolates. Tetracycline resistance (*tet(k); tet(38*)), β-lactam resistance *(mecA, blaZ);* Lincosamide resistance (*InuA);* Macrolide resistance *(erm(C); mph(C));* Aminoglycoside resistance *(ant(6)-la; aph(3')-lll);* Fusidic acid resistance *(fusB);* Streptogramin B resistance *(msr(A)*).

Isolates/ genes	*tet (k)*	*tet(38)*	*mecA*	*blaZ*	*lnu(A)*	*erm(C)*	*aph(3')-III*	*fusB*	*ant(6)-la*	*mph(C)*	*msr(A)*
MB46	-	+	+	+	+	+	+	+	-	-	-
MB6	+	+	+	+	-	-	+	+	+	-	-
MB7	-	+	+	-	-	-	-	-	-	-	-
MB9	-	+	+	+	-	+	+	+	+	-	-
MB12	+	+	+	+	-	-	+	+	-	-	-
CC80 reference strain	-	+	+	+	-	-	+	-	-	+	+

4.5. Plasmids content:

PlasmidFinder, revealed that all the *S. aureus* isolates carried plasmidic *rep* gene sequences. All isolates had *rep*7, MB46 only had *rep21*, while MB46 and MB9 had *rep10*. Finally, *rep 20* detected in all samples except in MB7 and the CC80 reference strain (table 1.5).

Table 5: Plasmids detected in the CC80 isolates and CC80 reference strain.

Isolates	*rep7*	*rep10*	*rep20*	*rep21*
MB46	+	+	+	+
MB6	+	-	+	-
MB7	+	-	-	-
MB9	+	+	+	-
MB12	+	-	+	-
Reference	+	-	-	-

4.6. Phage detection

At least two phage sequences were identified in each sample, with some sharing common phages such as phi7401PVL which was detected in MB12, MB9, MB6, MB46 and CC80 reference strain ST_1181-97 (Table 6).

Table 6 PHAST analysis showing the types of phages, length, position, completeness, score, #CDS (coding DNA sequence), and %GC for all studied isolates along with the CC80 reference strain.

Isolates	REGION	REGION_LENGTH	COMPLETENESS	SCORE	#CDS*	REGION_POSITION	POSSIBLE PHAGE	GC%
MB12	1	52.8Kb	intact	136	66	66-52918	PHAGE_Staphy_phi7401PVL_NC_020199	33.10%
	2	43.3Kb	intact	140	61	2206098-2249450	PHAGE_Staphy_JS01_NC_021773,	33.25%
MB9	1	10Kb	incomplete	30	8	1949300-1959299	PHAGE_Staphy_37_NC_007055	32.21%
	2	84.5Kb	intact	150	107	2216030-2300581	PHAGE_Staphy_phi7401PVL_NC_020199	33.06%
	3	38.1Kb	intact	110	47	2484057-2522241	PHAGE_Staphy_phiMR25_NC_010808	34.58%
	4	22.7Kb	incomplete	40	34	2823591-2846313	PHAGE_Staphy_JS01_NC_021773	30.35%
MB7	1	68.9Kb	intact	130	66	1802754-1871659	PHAGE_Staphy_phiMR25_NC_010808	34.10%
	2	95.6Kb	intact	150	125	2094540-2190222	PHAGE_Staphy_JS01_NC_021773	32.99%
MB6	1	42.1Kb	intact	110	62	2132413-2174540	PHAGE_Staphy_JS01_NC_021773	33.31%
	2	61.2Kb	intact	129	61	2610760-2672014	PHAGE_Staphy_phi7401PVL_NC_020199	34.05%
MB46	1	62.1Kb	intact	150	78	789713-851812	PHAGE_Staphy_phi7401PVL_NC_020199	32.74%
	2	42.7Kb	intact	130	60	2396296-2439035	PHAGE_Staphy_JS01_NC_021773	33.21%
CC80 reference Strain	1	44.6Kb	intact	140	63	642451-687086	PHAGE_Staphy_SA12_NC_021801	34.30%
	2	62.4Kb	intact	150	76	1551827-1614319	PHAGE_Staphy_phi7401PVL_NC_020199	32.70%
	3	49.8Kb	intact	140	63	2066729-2116572	PHAGE_Staphy_JS01_NC_021773	32.90%

4.7. CRISPRs

Most prokaryotic genomes have structures called CRISPRs. They are made up of 25-50 base pair repeats set apart by specific spacer sequences of comparable length, which come from prior virus exposures. All isolates contained several CRISPR sequences in their genome. The position and length of these sequences were identified in each isolate (Table 7).

Table 7: CRISPRFinder analysis in the different isolates showing the start and end positions of CRISPRs and their length.

Isolates	Start position	End position	Length (bp)	Sample	Start position	End position	Length (bp)
MB46	483223	483307	84	**MB12**	344215	344299	84
	965100	965185	85		652365	652508	143
	1497102	1497183	81		1015415	1015496	81
	1963532	1963629	97		1481879	1481976	97
	2480350	2480494	144		2290765	2290909	144
	2616280	2616366	86		2301689	2301773	84
					2426811	2426897	86
MB9	565436	565521	85	**MB7**	124741	124826	85
	1097727	1097808	81		999074	999155	81
	1918314	1918407	93		1211058	1211151	93
	1951327	1951424	97		1244071	1244168	97
	2335417	2335561	144		1525174	1525258	84
	2346341	2346425	84		2225058	2225202	144
	2471465	2471551	86		2235982	2236066	84
					2361134	2361220	86
MB6	164194	164287	93				
	197207	197304	97				
	906002	906080	78				
	990666	990747	81				
	1159798	1159891	93				
	1386256	1386341	85				
	1805918	1806002	84				
	2080090	2080174	84				
	2090954	2091098	144				
	2272680	2272766	86				

4.8. Circular genomes comparison and visualization

The visualization of each isolate in terms of coding DNA sequence, rRNA, tRNA and phages was done through DNAPLOTTER (Figure 2). Also, each isolate was compared to the CC80 reference strain and results were generated by CGview (Figure 3). Only isolate MB46 was chosen as a representative and shown graphically in this section, the rest of the circular genome displays can be found in Annex I.

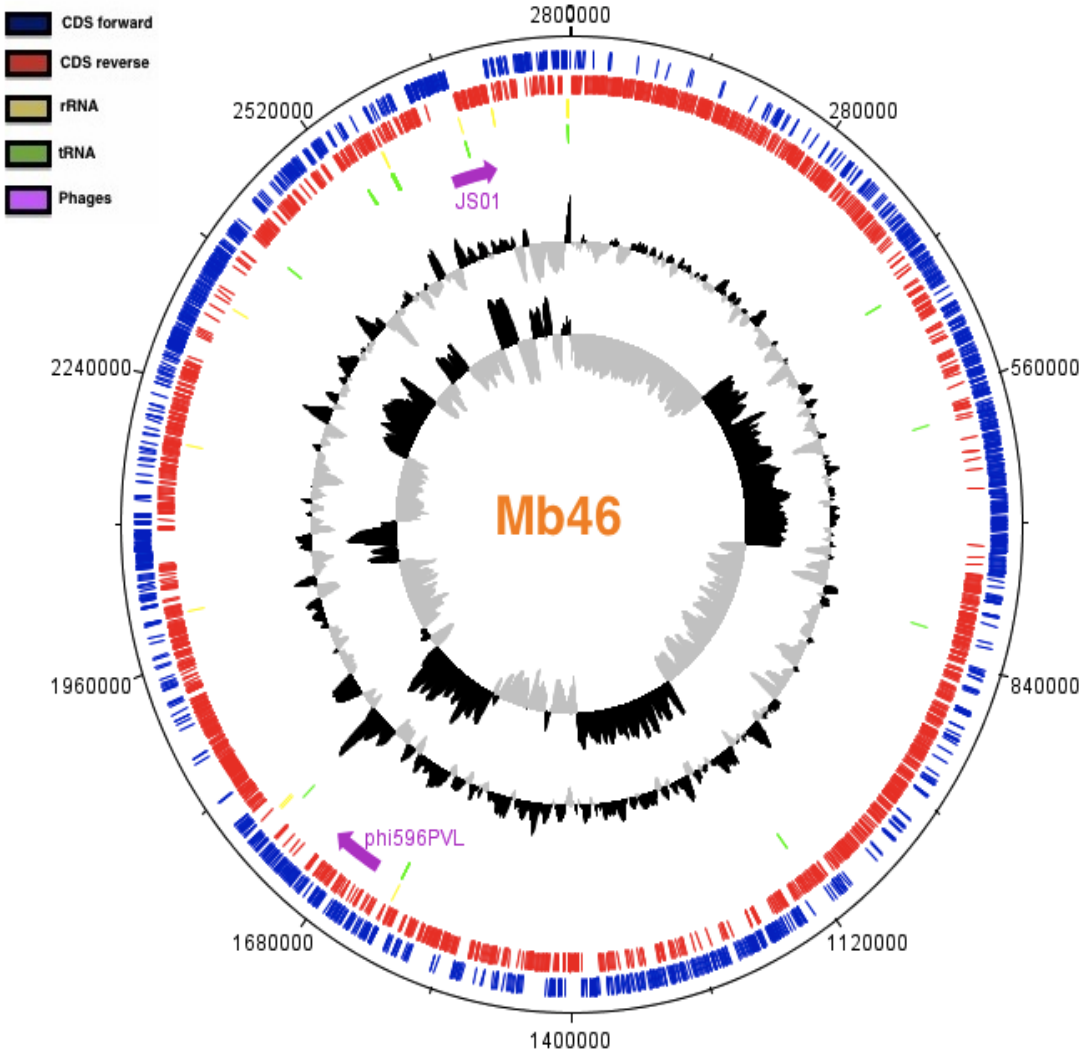

Figure 2: Genome atlas of *S. aureus* MB46. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand (blue); *S. aureus* coding DNA sequence on the reverse strand (red); ribosomal RNA genes (yellow); tRNA genes (green_ GC% (Black below mean and grey above mean); GC skew. The purple arrows represents the phages detected in the genome.

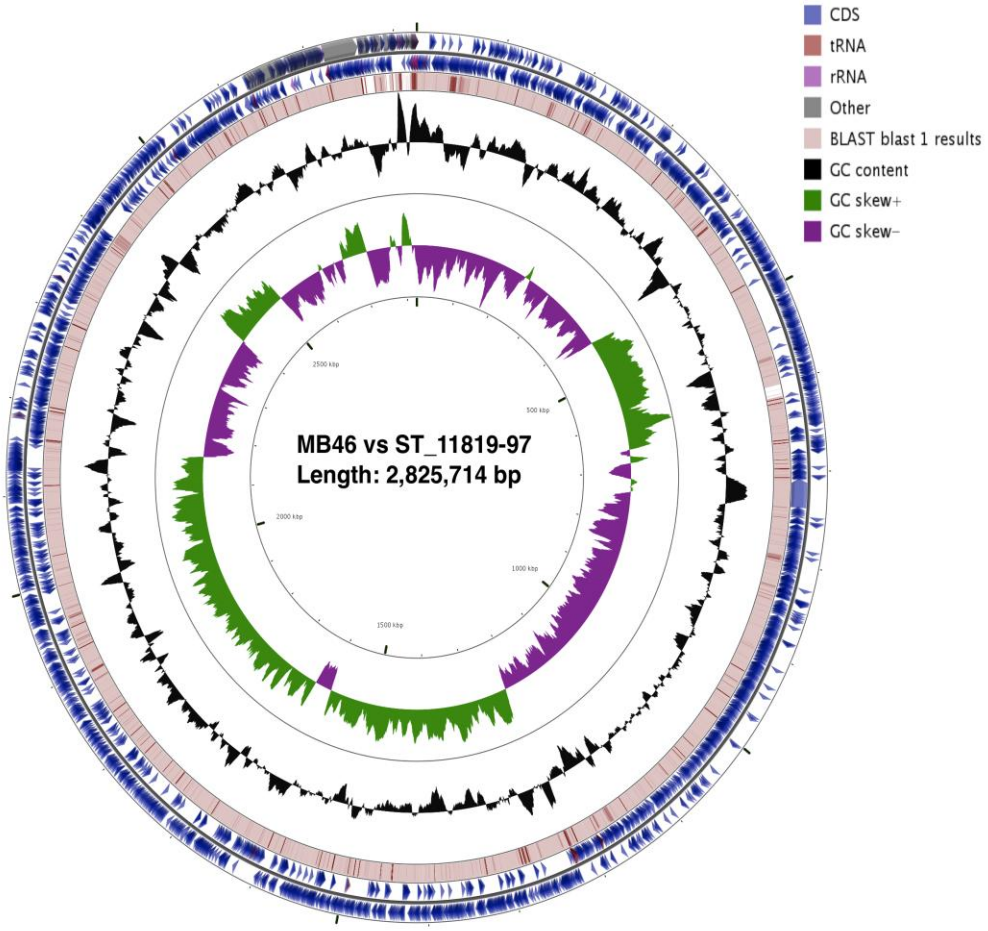

Figure 3: Genomic comparison between *S. aureus* MB46 and ST80 reference strain ST_11819-97. This graphical representation of the genome was generated using CGview server. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand, *S. aureus* coding DNA sequence on the reverse strand, rRNA and tRNA are colored according to the legend inside the first two tracks, third track, which is the blast track (Reference strain 11819_97), represents the positions covered by the BLASTN alignment. Inside this track, white regions indicate parts of the input sequence that did not yield a blast hit, light pink represents parts of the input sequence that yield one blast hit, darker pink regions indicate parts of the input sequence that yield several blast hits

(ie overlapping hits). These often include rRNA or tRNA genes or repetitive sequences for which several similar sequences in the CC80 reference strain are found. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively.

4.9 Phylogenetic tree.

A maximum parsimony phylogenetic tree was generated by CSIPhylogeny based on high quality SNPs and manipulated using FigTree v1.4.2. The branches of our CC80 isolates along with the reference strain (ST 11819-97) and a Dutch CC80 MRSA were colored by yellow. The remaining branches represent the clustering of the different clonal complexes (CC30, CC398, CC45, CC59, CC93, CC5, CC8 and CC1) and were colored in blue.

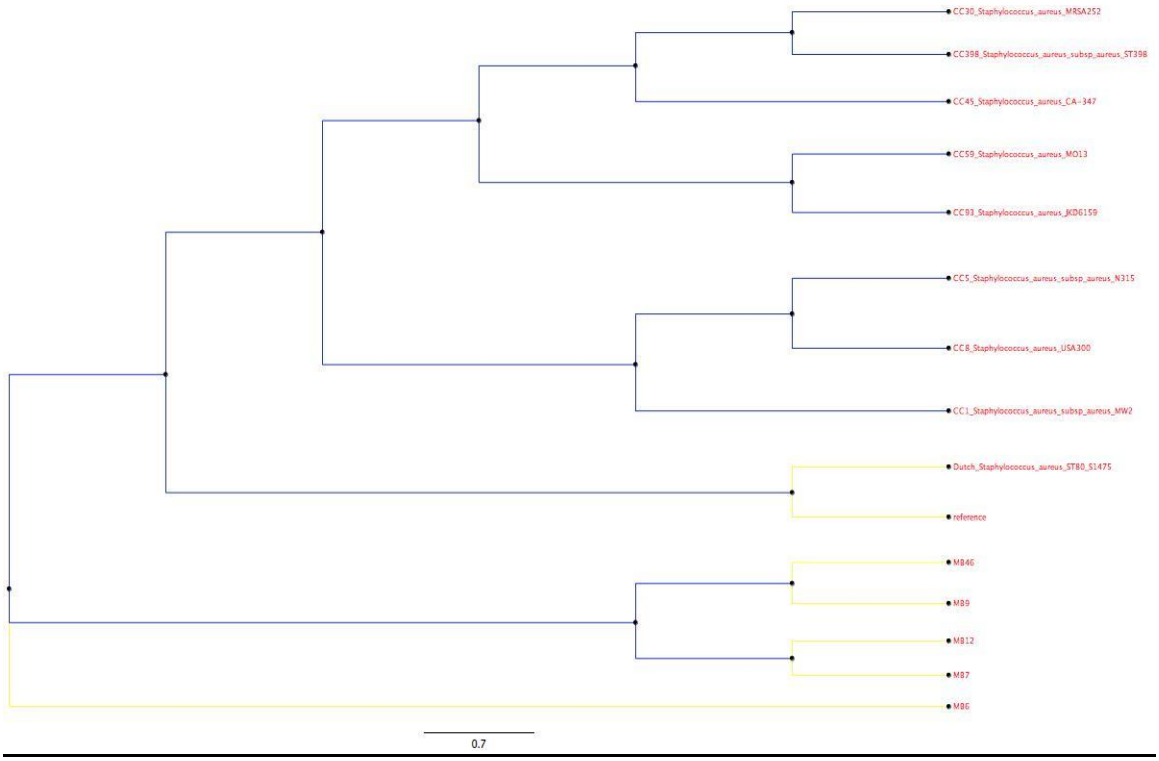

Figure 4 An urooted maximum parsimony SNP-based phylogenetic tree including the five isolates along with ten reference strains of different clonal complexes.

Chapter Five

DISCUSSION

5.1. PVL:

All the isolates in this study were found to be PVL positive. *S. aureus* strains encoding the PVL toxin, a prophage-encoded bicomponent pore-forming protein, are highly pathogenic and are associated with primary skin infections and life-threatening invasive diseases (Shrestha et al., 2014). PVL is a β-pore-forming toxin regarded as a cytotoxin (Pervost et al., 1995), which is mainly detected in CA-MRSA (Szmiegielsk et al., 1999; Kaneko et a., 2004). Phi7401PVL phage was detected in the genome of three of our isolates (MB46, MB6, MB12) along with the reference strain (ST_11819-7). This phage carried the PVL encoding genes precursors in all the former isolates except MB6. This might be the outcome of a deletion event that has occurred within the phage genome before being integrated. Our findings are similar to those reported by Mariem et al., (2013) where FG80 *S. aureus* isolates became more pathogenic upon acquiring phi7401PVL (Mariem et al., 2013). It's noteworthy however, that different types of phages can carry the PVL encoding genes. The majority of the isolates that were studied to track the evolution of CA-MRSA in Europe, were found to carry the PVL encoding genes on phiSA2, which was not detected in any of our isolates (Stegger et al., 2014).
In Lebanon a high percentage of MRSA were found to be PVL positive (Tokajian et al., 2010; Harastani et al., 2014). Interestingly, and in Denmark, the PVL positive CC80-ST80-IV isolates are the predominant cause of CA-MRSA infections with a large majority of patients having family relationships in the Middle East (Goering et al., 2009).

5.2. Biofilm Formation

Biofilm formation enables *S. aureus* to persist and to resist the host immune response by shielding them from opsonophagocytosis and antimicrobial agents (Foster, 2005). Biofilm formation requires two consecutive steps: primarily the adhesion of cells to a concrete substrate, followed by its adhesion to nearby cells creating multiple layers of bacteria. Cramton et al. (1999), showed that isolates having the ability to form biofilm carried the *ica* operon, which was also detected in all the isolates undertaken in this study. *ica* operon (*icaA, icaB, icaC* and *icaD*) mediates the production of polysaccharide intercellular adhesion (PIA). PIA is mainly composed of N –acetylglucosamine used for biofilm formation and invasion (Frank et al., 2011). Additionally, *ica* is needed for the expression of two other virulence determinants namely the genes encoding for extracellular matrix proteins (EMP) and EAP. Genes encoding both proteins were detected in all the studied isolates, which further demonstrated their pathogenicity. Interestingly, In most isolates (MB46, MB6, MB7 and MB12) the EAP gene was detected in close proximity with the JS01 phage sequence. Moreover, EAP was recently reported as being an important factor involved in inhibiting neutrophil serine proteases, and thus promoting *S. aureus* infection (Johnson et al., 2014). This finding increases our understanding into the complexity of *S. aureus* pathogenesis.

Additionally, the ability of *S. aureus* to colonize was linked to several surface proteins having high affinity to bind to MSCRAMMs. There proteins recognize fibronectin (Fn) such as FnBPA and FnBPB and are known to be essential for primary attachment (McCourt et al., 2014). Genome annotation of the isolates in this study revealed the presence of FnBPA and FnBPB. FnBPA is a protein that is found on the surface of *S. aureus* (Jönsson et al., 1991), and it attaches to both fibrinogen and Fn. This protein acts in mediating adhesion and invasion (Sinha et al., 2000; Saddif et al., 2001), and disrupting the genes coding for FnBP was associated with reduced invasiveness and adherence (Sinha et al., 2000).

Furthermore, FnBPs mediate the attachment of this pathogen to Fn coated biomaterials and were reported as being important factors enhancing internalization in experimental endocarditis (Menzies et al., 2003). On the other hand, Saddif et al. (2001) revealed that

the uptake of *S. aureus* by osteoblasts was linked to FnBPs. Disrupting the genes coding for these proteins blocked internalization, which further supports the essentiality of these factors as important virulent determinants.

5.3. Clumping factors

ClfA and ClfB, both of which were detected in the sequenced isolates, are clumping factors found on the surface of *S. aureus*. ClfA has high binding affinity to fibrinogen, which is a precursor of fibrin (Hawiger et al., 1982). Fibrin is one of the foremost components of plasma, which could be additionally found on biomedical materials (Vaudaux et al., 1989; 1995; Herrmann et al., 1993; Moreillon et al., 1995). Both ClfA and ClfB are members of the MSCRAMMs family (McDevitt et al., 1994; Ní Eidhin et al., 1998) sharing common characteristics with other cell surface proteins, but are characterized by their R region which is made up of serine and aspartic acid dipeptides (Josefsson et al., 1998). Several studies conducted on human platelets showed the role of ClfA and ClfB in aggregation (Brien et al., 2002). ClfB also mediates the adhesion of *S. aureus* to type I cytokeratin enhancing nasal colonization (Brien et al., 2002). Genome annotation of the isolates revealed the presence of both clumping factors. Similarly, previous studies also reported the presence of both factors in different subtypes including: ST8, ST5, ST256, ST88, ST188, ST552. On the other hand, and in contrast to our results, other ST80 isolates had only ClfA (McDevitt et al., 1994; Moreillon et al., 1995; Vaudaux et al., 1995). This illustrates that although the isolates undertaken in this study were typed as ST80, but were not exactly identical to other similar subtypes isolated from elsewhere, and thus studying such factors would be particularly important when testing for vaccines (McCarthy et al., 2010).

5.4. Exfoliative toxin

S. aureus produces different types of toxins including enterotoxins, exfoliative toxins and toxic shock syndrome toxin-1 (TSST-1) (Dinges et al., 2000). Exfoliative toxins have a serine protease domain, which acts on desmosomal cadherins leading to SSSS

(Shamez, 2003). SSSS is associated with two exfoliative toxin coding genes, namely *eta* and *etb,* the expression of which is regulated by *agr*. *eta* was commonly detected in isolates from Africa and USA, while *etb* was dominant in Japan (Bukowski et al., 2010). Both toxins were detected in MSSA and MRSA at different frequencies, with a 3-4% of MSSA having either *eta* or *etb* and 10% MRSA carrying *eta.* (Megevand et al., 2010; Silia et al., 2009). Moreover, *etd*, which is another toxin in this group, was detected in strains identified as belonging to CC80 (Holmes et al., 2005).

In this study all five isolates had *eta* and/or *etb* encoding genes. Likewise, in China, studying 108 *S. aureus* isolates taken from different hospitals had the *eta* gene within different clonal complexes such as CC5, CC7, CC25 and CC59 (Xie et al., 2011). However, *eta* and *etb* were both detected in *S. aureus* recovered from Taiwanese children with SSSS or staphylococcal toxic shock syndrome (STSS) (Chi et al., 2006). It's noteworthy that all the isolates in this study, had both the *eta* and PVL encoding genes, which was in harmony with previous results from Lebanon (Harastani et al., 2014). Contrary to the CC80 isolates, Kurt et al. (2013) found that PVL and exfoliative toxins were mutually exclusive in CC121 isolates.

5.5. Resistance patterns:

Differences in the antibiotic regimen that exists between countries contributed to the development of heterogeneous strains. CA-MRSA lineages are genotypically and phenotypically unrelated to the multi-drug resistant HA-MRSA. Common within CA-MRSA is that fact that they are generally susceptible to non-β-lactam antibiotics (Deresinski et al., 2005). The mechanisms underlying the unique SKF (Strepromycin, Kanamycin, Fusidic acid) resistance profile previously reported within the CC80 strains collected from Jordan and Lebanon, were revealed upon the annotation of the isolates (Harastani, et a., 2014). Resistance to Streptomycin was found to be through the *ant(6)-Ia* gene, which codes for an O-adenyltransferase that inactivates the drug through catalyzing the ATP-dependent adenylation of the drug (Shaw et al., 1993). Additionally, *aph(3')-111* gene which codes for O-phosphotransferase that renders the bacterium to be resistant to vancomycin, Neomycin, Paromomycin, Kanamycin and Gentamycin B was

also detected. Finally, resistance to fusidic acid is another common important characteristic feature of CC80 strains, and which was previously observed and reported by several studies conducted in the region and elsewhere (Tokajian et al, 2010; Harsatani et al., 2014, Stegger et al., 2014, Goering et al., 2009). Sequencing of the isolates revealed the presence of the *fusB* gene. The *fusB* gene was found to be located on phage-related resistance islands in *S. epidermidis* (Hsiao et al., 2011), and on p11819-97 plasmid in CC80 strains, which additionally carried the *blaZ* gene (Stegger et al., 2014).

Plasmids are classified based on the replication proteins (REP) (Jensen et al., 2010). Plasmids may carry a single *rep* gene or a combination of *rep* genes, with each combination of *rep* forming a plasmid group of *S. aureus (pGSA).* We have detected the presence of different *rep* genes associated with a set of other genes encoding for resistance. *rep7*, which was detected in all our isolates is classified as *pGSA2. pGSA2* is known to carry four variable resistance encoding genes being namely *cat* (chloramphenicol resistance)*, tetK (tetracycline resistance) str* (streptomycin resistance), *vgaA* (pleuromutilin resistance) (McCarthy et al., 2012). This however, matched with the annotation results were *tetK* was one of the detected genes. Moreover, *rep20* which was detected in MB46, MB6, MB9 and MB12, is classified as *pGSA31* and has *blaZ, cadA, tetK* as core genes, while *rep10* is *pGSA3* carrying only one core gene which is *ermC*, and was also detected in this study upon annotation. Finally, *rep21*, which is *pGSA10,* carries variable resistance genes including *blaZ*, which was also detected upon annotation.

5.7. CRISPR identification:

Most prokaryotic genomes have structures called CRISPRs. They are made up of 25-50 base pair repeats set apart by specific spacer sequences of comparable length, which come from prior virus exposures (Marrafini et al., 2010). *cas* genes and CRISPRs are found in the genomes of archaea (≈90%) and bacteria (≈40%) (Kunin et al*.*, 2007*;* Grissa et al., 2007) and they have been proven to play a role in sequence-targeted immunity

towards plasmids (Marraffini et al., 2008) and phages (Brouns et al., 2008; Barrangou et al., 2007). This system could be responsible for the relatively lower number of phage-associated virulence factors and antibiotic resistance genes detected in ST80 MRSA in comparison to other widespread MRSA strains (Monecke et al., 2008; Walther et al., 2009). Furthermore, it has been proposed that the transfer of CRISPRs occurs via horizontal gene transfer by mega-plasmid (greater than 40 kb) conjugation, with their origin remaining unfamiliar (Godde et al., 2006). Based on CRISPRfinder, the studied isolates had CRISPRs sequences at different positions within their genomes. However, only one CRISPR system identified in MB9, was detected within a putative mobile genetic element that extends from base pair number 1951327 to 1951424. Accordingly, and since CRISPR was detected on a mobile element, we propose and in harmony with Golding et al. (2010), that additional mechanism of mobilization to other *Staphylococcus spp*. is plausible (Golding et al., 2010). CRISPRs are widely being studied nowadays in terms of developing new therapeutic approaches. Antimicrobial agents are uncapable of selectively killing specific members within a mixed microbial population since they are designed to target growth functions or conserved prokaryotic cellular pathways. Therefore, sequence specific and programmable antibiotics are being developed using Cas9 (Jinek et al., 2012; Gasiunas et al., 2012), and through introducing an RNA-guided nuclease by bacteriophages. When the nuclease is reprogrammed to target genes coding for antibiotic resistance in staphylococci, it will degrade plasmids carrying antibiotic resistance genes, and consequently reduce the transfer of plasmid-borne resistance genes to avirulent staphylococci (Diep et al., 2006; Weigel et al., 2003). Reprogrammed Cas9 directed towards virulence genes could help in targeting pathogenic strains of *S. aureus (*Bikard et al., 2014).

5.8. Phylogeny

Phylogenetic examination determined *S. aureus* MW2 to be the out-group most closely related to our CC80 isolates when compared to the remaining genome sequences of different clonal complexes, and this was similar to what was previously reported (Stegger et al., 2014).

All our isolates, along with the reference strain (ST_18819-97) and the Dutch ST80 (S1475), were clustered together. This is predictable since they all belong to the CC80 and share common characteristics in terms of virulence, resistance and other genomic features. Only MB6 was distantly placed, this can be attributed to differences seen in its resistance profile and to the high number of SNPs detected when aligned against the CC80 reference strain.

5.9. Importance of this study

To summarize, the WGS of each isolate helped us in roughly characterizing most of the genome and identifying several genes involved in virulence, resistance and many other sequences corresponding to different types of mobile elements. PVL was detected along with several factors involved in biofilm formation, invasion and toxin production. Also different resistance profiles were linked to the underlying resistance mechanisms. This can improve the appropriate use of antimicrobials by promoting the selection of the optimal antimicrobial drug regimen to achieve best clinical outcome. Finally, WGS cuts the time needed to diagnose infections, improving patient care, optimizing clinical outcome and enabling more precise and appropriate individualized therapies, improving patient care and optimizing clinical outcome.

Chapter Six

CONCLUSION

- In this study, which is the first in Lebanon, genome sequencing of CC80 *S. aureus* isolates from Lebanon and Jordan allowed us to characterize our samples in terms of virulence, resistance, mobile element, CRISPRs and phylogeny

- All our isolates were PVL positive and carried genes encoding for factors involved in biofilm formation such as *icaA, icaB, icaC*, *icaD,* EMP and EAP.

- Fnb were present in all our isolates illustrating invasiveness.

- All the isolates had clumping factors A and B, both of which are linked to adhesion.

- The *eta* gene coding for exfoliative toxin A and which plays a role in Staphylococcal Scalded Skin Syndrome was also detected.

- Different resistance patterns of all isolates were detected, with all being tetracycline and β-lactam resistant, MB46 only lincosamide, streptogramin B and macrolide resistant, and both MB9 and CC80 reference strain were also resistant to macrolide. All isolates except MB7 were aminoglycoside resistant, and three isolates (MB46, MB6 and MB9) were fusidic acid resistant

- Plasmids were identified in all our isolates and were correlated to the resistance genes found in each.

- Integrated phages were detected in all our isolates showing the potential of resistance gene transfer.

- CRISPR sequences along with *cas* were identified. The latter shows a great promise for future medicinal use targeting multidrug resistant strains.

- Phylogenetic examination determined *S. aureus* MW2 to be the out-group most closely related to our CC80 isolates when compared to the remaining genome sequences of different clonal complexes.

BIBLIOGRAPHY

Abdelnour, A., Arvidson, S., Bremell, T., Ryden, C., & Tarkowski, A. (1993). The accessory gene regulator (agr) controls *Staphylococcus aureus* virulence in a murinearthritis model. *Infection Immunity*, *61*(9), 3879–3885.

Ahmed, S., Meghji, S., Williams, R. J., Henderson, B., Brock, J. M., & Nair, P. S. (2001). *Staphylococcus aureus* fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria. *Infection and Immunity*, *69*(5), 2872-2877.

Álvarez, G. L., Holden, M. T., Linsday, H., Webb, C. R., Brown, D. F., Curran, M. D., … Holmes, M. A. (2011). Meticillin-resistant *Staphylococcus aureus* with a novel *mecA* homologue in human and bovine populations in the UK and Denmark: A descriptive study. *The Lancet Infectious Diseases*, *11*(8), 595-603.

American Academy of Orthopedics Surgeons. (n. d.) The silent epidemic : CA-MRSA and HA-MRSA. Retrieved March 10, 2015, from http://www.aaos.org/news/aaosnow/may08/research1.asp

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., … Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, *9*, 75. doi:10.1186,1471-2164-9-75

Baba, T. (2014). Overview of bacterial genome analysis, and genomic diversity and virulence of *Staphylococcus aureus. Jutendo Medical Journal, 60*(1), 63-65.

Bachert, C., Gevaert, P., & Cauwenberge, V. P. (2002). *Staphylococcus aureus* enterotoxins: A key in airway disease? *Allergy*, *57*(6), 480-487.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., … Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5829), 1709–1712.

Bikard, D., Euler, C. W., Jiang, W., Nussenzweig, P. M., Goldberg ,G. W., Duportet, X., … Marraffini, L. A. (2014). Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nature Biotechnology*, *32*, 1146-1150.

Bradley, S. F. (2005). *Staphylococcus aureus* pneumonia: Emergence of MRSA in the community. *Seminar in Respiratory and Critical Care Medicine* , *26*(6), 643–649.

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., … vanderOost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes, *Science, 321*(5891), 960-964.

Bukowski, M., Wladyka, B., Dubin, G. (2010). Exfoliative toxins of *Staphylococcus aureus. Toxins, 2*(5), 1148-1165.

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Volby Larsen, M., Lund, Villa, L., Møller-Aarestrup, F. & Hasman, H. (2014). Plasmid finder and pMLST: In silico detection and typing of plasmids. *Antimicrobial Agents and Chemotherapy, 58*(7), 3895-903. doi:10.1128/AAC.02412-14

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics*, *25*(1), 19-20.

Chambers, H. F., & Deleo, F. R., (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Review. Microbiology*, *7*(9), 629–641.

Chang, S., Sievert, D. M., Hageman, J. C., Boulton, M. L., Tenover, F. C., Downes, F. P., … Fridkin, S. K. (2003). Infection with vancomycin-resistant *Staphylococcus aureus* containing the *vanA* resistance gene. *The New England Journal of Medicine, 348*(14), 1342–1347.

Chevreux, B., Wetter, T., & Suhai, S. (1999). Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology. *Proceedings of the German Conference on Bioinformatics*, *99*, 45-56.

Chi, C. Y., Wang, S. M., Lin, H. C., & Liu, C. C. (2006). A clinical and microbiological comparison of *Staphylococcus aureus* toxic shock and scalded skin syndromes in children. *Clinical Infectious Diseases, 42*(2), 181-185.

Clauditz, A., Resch, A., Wieland, K. P., Peschel, A., & Gotz, F. (2006). Staphyloxanthin plays a role in the fitness of *Staphylococcus aureus* and its ability to cope with oxidative stress. *Infection and Immunity, 74*(8), 4950-4953.

Cosentino, S., Larsen, M. V., Aarestrup, F. M., & Lund, O. (2013). PathogenFinder - distinguishing friend from foe using bacterial whole genome sequence data. *PLOS One, 8*(10), e77302. doi: 10.1371/journal.pone.0077302

Cramton, S. E., Gerke, C., Schnell, N. F., Nichols, W. W., & Götz, F. (1999). The intercellular adhesion (ica) locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infection and Immunity*, *67*(10), 5427-5433.

DeDent, A. C., McAdow, M., & Schneewind, O. A. (2007). Distribution of protein A on the surface of *Staphylococcus aureus. Journal of Bacteriology*, *189*(12), 4473-4484.

Deresinski, S. (2005). Methicillin-resistant *Staphylococcus aureus*: An evolutionary, epidemiologic and therapeutic odyssey. *Clinical Infectious Diseases, 40*(4), 562–573.

Deurenberg, R. H., & Stobberingh, E. E. (2008). The evolution of *Staphylococcus aureus. Infection Genetics and Evolution, 8*(6), 747-763.

Diep, B. A., Gill, S. R., Chang, R. F., Phan, T. H., Chen, J. H., Davidson, M. G., … Perdreau-Remington, F. (2006). Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus*. *Lancet*, *367*(9512), 731–739.

Dinges, M. M., Orwin, P. M., & Schlievert, P. M. (2000). Exotoxins of *Staphylococcus aureus*. *Clinical Microbiology Review, 13*(1), 16–34.

Dryden, M. S. (2010). Complicated skin and soft tissue infection. *Journal of Antimicrobial Chemotherapy*, *65*(suppl 3), 35-44.

Enright, M. C., Robinson, D. A., Randle, G., Feil, E. J., Grundmann, H. & Spratt, B. G. (2002). The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proceedings of the National Academy of Sciences of United States of America*, *99*(11), 7687-7692.

Foster, T. J. (2005). Immune evasion by staphylococci. *Nature Reviews Microbiology*, *3*(12), 948-958

.

Frank, K. L., & Patel, R. (2007). Poly-N-acetylglucosamine is not a major component of the extracellular matrix in biofilms formed by icaADBC-positive *Staphylococcus lugdunensis* isolates. *Infection and Immunity, 75*(10), 4728-4742.

Fricke, F., & Rasko, D. (2013). Bacterial genome sequencing in the clinic: Bioinformatic challenges and solutions. *Nature Reviews Genetics, 15*, 49-55.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceeding of the National Academy of Sciences in the United States of America, 109*(39), 2579-2586.

Genestier, A. L., Michallet, M. C., Prevost, G., Bellot, G., Chalabreysse, L., Peyrol, S., … Genestier, L. (2005). *Staphylococcus aureus* Panton-Valentine Leukocidin directly targets mitochondria and induces Bax-independent apoptosis of human neutrophils. *The Journal of Clincal Investigation*, *115*(11), 3117–3127.

Godde, J. S., & Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution, 62*(6), 718–729.

Goering, R. V., Larsen, A. R., Skov, R., Tenover, F. C., Anderson, K. L., & Dunman, P. M. (2009). Comparative genomic analysis of European and Middle Eastern community-associated methicillin-resistant *Staphylococcus aureus* (CC80:ST80-IV) isolates by high-density microarray. *Clinical Microbiology and Infection, 15*(8), 748-755.

Golding, G. R., Bryden, L., Levett, P. N., McDonald, R. R., Wong, A., Graham, M. R., … Mulbey, M. R. (2010). Whole-genome sequence of livestock-associated st398 methicillin-resistant *Staphylococcus aureus* isolated from humans in Canada. *Journal of Bacteriology, 194*(23), 6627-6628.

Gordon, R. J., & Lowy, F. D. (2008). Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clinical Infectious Disease, 46*(Suppl 5), 350-359.

Gould, I. M. (2010). VRSA-doomsday superbug or damp squib? *Lancet Infectious Diseases, 10*(12), 816–818.

Graves, S. F., Kobayashi, S. D., & DeLeo, F. R. (2010). Community-associated methicillin-resistant *Staphylococcus aureus* immune evasion and virulence. *Journal of Molecular Medicine, 88*(2), 109-114.

Gregg, A. R., Van den Veyver, I. B., Gross, S. J, Madankumar, R., Rink, B. D., & Norton, M. E. (2014). Noninvasive prenatal screening by next-generation sequencing. *Annual Review of Genomics and Human Genetics, 15*, 327-347.

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, *35*, W52-7.

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, *8*(172), 1471-2105.

Hammer, N. D., & Skaar, E. P. (2011). Molecular mechanisms of *Staphylococcus aureus* iron acquisition. *Annual Review of Microbiology*, *65*, 129-147.

Hanaki, H., Kuwahara-Arai, K., Boyle-Vavra, S., Daum, R. S., Labischinski, H., & Hiramatsu, K. (1998). Activated cell-wall synthesis is associated with vancomycin resistance in methicillin-resistant *Staphylococcus aureus* clinical strain MU3 and MU50. *The Journal of Antimicrobial Chemotherapy*, *42*(2), 199-209.

Hanaki, H., Labischinski, H., Inaba, Y., Kondo, N., Murakami, H., & Hiramatsu, K. (1998). Increase in glutamine-non-amidated muropeptides in the peptidoglycan of vancomycin-resistant *Staphylococcus aureus* strain Mu50. *The Journal of Antimicrobial Chemotherapy*, *42*(3), 315–320.

Harastani, H. H., & Tokajian, S. T. (2014). Community-associated methicillin-resistant *Staphylococcus aureus* clonal complex 80 type IV (CC80-MRSA-IV) isolated from the Middle East: A heterogeneous expanding clonal lineage. *Plos One, 9*(7), e103715.

Harris, S. R., Cartwright, E. J. P., Török, M. E., Holden, M. T. G., Brown, N. M., Ogilvy-Stuart, A. L., … Peacock, S. J. (2013). Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: A descriptive study. *The Lancet Infectious Diseases, 13*(2), 130-136.

Hartman, B. J., & Tomasz, A. (1984). Low-affinity penicillin-binding protein associated with beta-lactam resistance in Staphylococcus aureus, *Journal of Bacteriology*, *158*(2), 513–516.

Hawiger, J., Timmons, S., Strong, D. D., Cottrell, C. A., Riley, M., & Doolittle, R. F. (1982). Identification of a region of human fibrinogen interacting with staphylococcal clumping factor. *Biochemistry*, *21*(6), 1407–1413.

Herrmann, M., Lai, Q. J., Albrecht, R. M., Mosher, D. F., & Proctor, R. A. (1993). Adhesion of *Staphylococcus aureus* to surface-bound platelets: Role of fibrinogen/fibrin and platelet integrins. *The Journal of Infectious Diseases, 167*(2), 312–322.

Holmes, A., Ganner, A., McGcuane, S., Pitt, T. L., Cookson B. D., & Kearns, A. M. (2005). Staphylococcus aureus isolates carrying Panton-Valentine Leucocidin genes in England and Wales: Frequency, characterization, and association with clinical disease**.** *Journal of Clinical Microbiology, 43*(5), 2384-2390.

Hsiao, H., Chuang, C., Tan, Y., Ma, H., Lin, K., & Chang, J. (2011). Methicillin-resistant *Staphylococcus aureus* ocular infection, a 10-year hospital-based study. *Ophthalmology, 199*(3), 522-527.

Jacoby, G. A. (2005). Mechanisms of resistance to quinolones. *Clinical Infectious Diseases, 41*(suppl 2), 120-126.

Jevons, M. P. (1961). Celbenin-resistant staphylococci. *British Medical Journal*, *2*(5212), 124–125.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2009). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337*(6096), 816–821.

Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M., & Aarestrup, F.M. (2014). FM real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli. Journal of Clinical Microbiology, 52*(5), 1501-1510.

Johannessen, M., Sollid, J. E., & Hanssen, A. (2012). Host and microbe determinants that may influence the success of *S. aureus* colonization. *Frontiers in Cellular and Infection Microbiology*, *2*(56), 1-14.

Johnson, M., Cockayne, A. & Morrissey, J. A. (2008). Iron-regulated biofilm formation in *Staphylococcus aureus* Newman requires ica and the secreted protein Emp. *Infection and Immunity*, *76*(4), 1756-1765.

Jönsson, K., Signäs, C., Müller, H. P., & Lindberg, M. (1991). Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*: The complete nucleotide sequence and characterization of the second gene. *European Journal of Biochemistry, 202*(3), 1041-1048.

Josefsson, E., McCrea, K. W., Ní Eidhin, D., O'Connell, D., Cox, J., Höök, M., & Foster, T. J. (1998). Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus. Microbiology, 144*(Pt12), 3387-3395.

Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., & Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLOS One*, *9*(8), e104984.

Kaneko, J., & Kamio, Y. (2004). Bacterial two-component and hetero-heptameric pore-forming cytolytic toxins: Structures, pore-forming mechanism, and organization of the genes. *Bioscience and Biotechnology Biochemistry, 68*(5), 981–1003.

Kim, S. H., Park, W. B., Lee, K. D., Kang, C. L., Kim, H. B., Oh, M. D., … Choe, K. W. (2003). Outcome of *Staphylococcus aureus* bacteremia in patients with eradicable foci versus noneradicable foci. *Clincal Infectous Disease*, *37*(6), 794–799.

Kirst, H. A., Thompson, D. G., & Nicas, T. I. (1998). Historical yearly usage of vancomycin. *Antimicrobial Agents and Chemotherapy*, *42*(5), 1303–1304.

Kluytmans, J., Van Belkum, A., & Verbrugh, H. (1997). Nasal carriage of Staphylococcus aureus: Epidemiology, underlying mechanisms, and associated risks. *Clinical Microbiology Reviews, 10*(3), 505-520.

Koboldt, D., Meltz Steiberg, K., Larson, D., Wilson, R., & Mardis, E. (2013). The next-generation sequencing revolution and its impact on genomics. *Elsevier*, *155*(1), 27-38.

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., Holden, M. T. G., … & Peacock, S. J. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathogens*, *8*(8), e1002824.

Kunin, V., Sorek, R., & Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biology*, *8*(4), R61.

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., … Hiramatsu, K. (2001). Whole genome sequencing of meticillin-resistant *Staphylococcus aureus. Lancet, 357*(9264), 1225-1240.

Kurt, K., Rasigade, J. P., Laurent, F., Goering, R. V., Žemličková, H., Machova, I., … Nübel. (2013). Subpopulations of *Staphylococcus aureus* clonal complex 121 are associated with distinct clinical entities. *PLoS One, 8*(3), e58155.

LiU, G. Y. (2009). Molecular pathogenesis of *Staphylococcus aureus* infection. *Pediatric Research, 65*(5 Pt 2), 71R-77R.

Lowy, F. D. (2003). Antimicrobial resistance: The example of *Staphylococcus aureus*. *Science Medicine*, *111*(9), 1265-1273.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics, 9*, 387-402.

Mariem, B. J. J, Ito, T., Zhang, M., Jin, J., Li, S., Ilhem, B. B. B., … Hiramatsu, K. (2013). Molecular characterization of methicillin-resistant Panton-Valentine Leukocidin positive *Staphylococcus aureus* clones disseminating in Tunisian hospitals and in the community. *BMC Microbiology, 13*(2). doi: 10.1186/1471-2180-13-2

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science, 322*(5909), 1843–1855.

Marraffini, L. A., & Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics, 11*(3), 181–190.

McCarthy, A. J., & Lindsay, J. A. (2010). Genetic variation in *Staphylococcus aureus* surface and immune evasion genes is lineage associated: Implications for vaccine design and host-pathogen interactions. *BMC Microbiology, 10*(173). doi:10.1186/1471-2180-10-173

McCarthy, A. J., & Lindsay, J. A. (2012). The Distribution of plasmids that carry virulence and resistance genes in *Staphylococcus aureus* is lineage associated. *BMC Microbiology*, *12*(104). doi:10.1186/1471-2180-12-104

McCourt, J., O'Halloran, D. P., McCarthy, H., O'Gara, J. P., & Geoghegan, J. A. (2014). Fibronectin-binding proteins are required for biofilm formation by community-associated methicillin-resistant *Staphylococcus aureus* strain LAC. *FEMS Microbiology Letter, 353*(2), 157-164.

McDevitt, D., Francois, P., Vaudaux, P., & Foster, T. J. (1994). Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus. Molecular Microbiology*, *11*(2), 237–248.

Megevand, C., Gervaix, A., Heininger, U., Berger, C., Aebi, C., Vaudaux, B., … Francois, P. (2009). Molecular epidemiology of the nasal colonization by methicillin-susceptible *Staphylococcus aureus* in Swiss children. *Clinical Microbiology and Infection*, *16*(9), 1414-1420.

Menzies, B. E. (2003). The role of fibronectin binding proteins in the pathogenesis of *Staphylococcus aureus* infections. *Current Opinion in Infectious Diseases, 16*(3), 225-229.

Miller L. G., & Diep, B. A. (2008). Clinical practice: colonization, fomites, and virulence: rethinking the pathogenesis of community-associated methicillin-resistant *Staphylococcus aureus* infection. *Clinical Infectious Diseases, 46*(5), 752–760.

Monecke, S., Jatzwauk, L., Weber, S., Slickers, P., & Ehricht, R. (2008). DNA microarray-based genotyping of methicillin-resistant *Staphylococcus aureus* strains from eastern Saxony. *Clin Microbial Infection*, *14*(6), 534–545.

Moran, G. J., Krishnadasan, A., Gorwitz, R. J., Fosheim, G. E., McDougal, L. K., Carey, R. B., & Talan, D. A. (2006). Methicillin-resistant *S. aureus* infections among patients in the emergency department. *The New England Journal of Medicine*, *355*(7), 666–674.

Moreillon, P., Entenza, J. M., Francioli, P., McDevitt, D., Foster, T. J., Francois, P., & Vaudaux, P. (1995). Role of *Staphylococcus aureus* coagulase and clumping factor in pathogenesis of experimental endocarditis. *Infection Immunity*, *63*(12), 4738–4743.

Moroney, S. M., Heller, L. C., Arbuckle, J., Talavera, M., & Widen, R. J. (2007). staphylococcal cassette chromosome *mec* and Panton-Valentine Leukocidin characterization of methicillin-resistant *Staphylococcus aureus* clones. *Journal of Clinical Microbiology, 45*(3), 1019-1021.

Nair, D., Memmi, G., Hernandez, D., Bard, J., Beaume, M., Gill, S., … Cheung, A. L. (2011). Whole-genome sequencing of *Staphylococcus aureus* strain RN4220, a key laboratory strain used in virulence research, identifies mutations that affect not only virulence factors but also the fitness of the strain. *Journal of Bacteriology, 193*(9), 2332-2335.

Ní Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Höök, M., & Foster, T. J. (1998). Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus. Molecular Microbiology*, *30*(2), 245–257.

Ní Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Höök, M., & Foster, T. J. (1998). Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus. Molecular Microbiology*, *30*(2), 245–257.

O'Brien, L., Kerrigan, S. W., Kaw, G., Hogan, M., Penadés, J., Litt, D., … Cox, D. (2002). Multiple mechanisms for the activation of human platelet aggregation by *Staphylococcus aureus:* roles for the clumping factors ClfA and ClfB, the serine-aspartate repeat protein SdrE and protein A. *Molecular Biology, 44*(4), 1033-1044.

Odell, C. A. (2010). Community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) skin infections. *Current Opinion in Pediatrics, 22*(3), 273-277.

Ogston, A. (1889). Micrococcus poisoning. *Journal of Anatomy and Physiology, 17*(PT1), 24-58.

Ortega, E., Abriouel, H., Lucas, R., & Gálvez, A. (2010). Multiple roles of *Staphylococcus aureus* enterotoxins: Pathogenicity, superantigenic activity, and correlation to antibiotic resistance. *Toxins, 2*(8), 2117-2131.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T. … & Stevens, R. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Research, 42*, D206-D214.

Prevost. G., Cribier, B., Couppie, P., Petiau, P., Supersac, G., Finck-Barbancon, V., … Piemont, Y. (1995). Panton-Valentine Leucocidin and gammahemolysin from *Staphylococcus aureus* ATCC 49775 are encoded by distinct genetic loci and have different biological activities. *Infection and Immunity*, *63*(10), 4121-4129.

Rubinstein, E., Kollef, M. H., & Nathwani, D. (2008). Pneumonia caused by methicillin-resistant *Staphylococcus aureus*. *Clinical Infectious Diseases, 46*, 378-385.

Schlievert P. M., Strandberg K. L., Lin Y. C., Peterson M. L., & Leung D. Y. (2010). Secreted virulence factor comparison between methicillin-resistant and methicillin-sensitive *Staphylococcus aureus*, and its relevance to atopic dermatitis. *The Journal of Allergy and Clinical Immunology*, *125*(1), 39–49.

Shamez, L. (2003). Understanding the mechanism of action of the exfoliative toxins of *Staphylococcus aureus. FEMS Immunology and Medical Microbiology, 39*(2),181-189.

Shaw, K. J., Rather, P. N., Hare, R. S., & Miller, G. H. (1993). Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiological Reviews*, *57*(1), 138-163.

Shopsin, B., Drlica-Wagner, A., Mathema, B., Adhikari, R. P., Kreiswirth, B. N., & Novick, R. P. (2008). Prevalence of agr dysfunction among colonizing *Staphylococcus aureus* strains. *The Journal of Infectious Diseases, 198*(8), 1171–1174.

Showsh, S. A., De Boever, E. H., & Clewell, D. B. (2001). Vancomycin resistance plasmid in *Enterococcus faecalis* that encodes sensitivity to a sex pheromone also produced by *Staphylococcus aureus*. *Antimicrobial Agents Ch*emotherapy, *45*(7), 2177–2178.

Shrestha, B., Singh, W., Raj, V. S., Pokhrel, B. M., & Mohapatra, T. M. (2014). High prevalence of Panton-Valentine Leukocidin (PVL) genes in nosocomial-acquired *Staphylococcus aureus* isolated from tertiary care hospitals in Nepal. *BioMed Research International*, 2014, 790350. doi:10.1155/2014/790350

Sieprawska, M. L., Mydel, P., Krawczyk, K., Wo´jcik, K., Puklo, M., Lupa, B., … Potempa, J. (2004). Degradation of human antimicrobial peptide LL-37 by *Staphylococcus aureus* derived proteinases. *Antimicrobial Agents and Chemotherapy*, *48*(12), 4673-4679.

Sila, J., Sauer, P., & Kolar, M. (2009). Comparison of the prevalence of genes coding for enterotoxins, exfoliatins, Panton-Valentine Leukocidin and tsst-1 between methicillin-resistant and methicillin susceptible isolates of *Staphylococcus aureus* at the university hospital in Olomouc. *Biomedical Papers of Medicine Faculty University Palacky. Olomouc. Czech Repub*, *153*(3), 215–218.

Sinha, B., & Fraunholz, M. (2010). *Staphylococcus aureus* host cell invasion and post-invasion events. *International Journal of Microbiology*, *300*(2-3), 170-175.

Skinner, D., & Keefer, C. S. (1941). Significance of bacteremia caused by *Staphylococcus aureus. Archives of Internal Medicine, 68*(5), 851–875.

Smith, T. L., Pearson, M. L., Wilcox, K. R., Cruz, C., Lancaster, M. V., Robinson-Dunn, B., … Jarvis, W. R. (1999). Emergence of vancomycin resistance in *Staphylococcus aureus*: Glycopeptide-intermediate *Staphylococcus aureus* working group. *The New England Journal of Medicine, 340*(7), 493–501.

Stegger, M., Wirth, T., Andersen, P. S., Skov R. L., Grassi A. D., Simões, P. M., … Laurent, F. (2014). Origin and evolution of European community-acquired methicillin-resistant *Staphylococcus aureus. mBio, 5*(5), e01044-14.

Stothard, P., & Wishart, D. S. Circular genome visualization and exploration using CGView. *Bioinformatics*, *21*, 537-539.

Szmigielski, S., Prevost, G., Monteil, H., Colin, D. A., & Jeljaszewicz, J. (1999). Leukocidal toxins of staphylococci. *Zentralblatt Bakteriologie*, *289*(2), 185–201.

Tokajian, S. T., Khalil, P. A., Jabbour, D., Rizk, M., Farah, M. J., Hashwa, F. A., & Araj, G. F. (2010). Molecular characterization of *Staphylococcus aureus* in Lebanon. *Epidemiology and Infection*, *138*(5), 707–712.

Turleji, A., Hryniewiczi, W., & Empel, J. (2011). Staphylococcal cassette chromosome mec (SCC*mec*) classification and typing methods: An overview. *Polish Journal of Microbiology, 60*(2), 95-103.

Vaudaux, P., Pittet, D., Haeberli, A., Huggler, E., Nydegger, U. E., Lew, D. P., & Waldvogel, F. A. (1989). Host factors selectively increase staphylococcal adherence on inserted catheters: A role for fibronectin and fibrinogen or fibrin. *The Journal of Infectious Diseases*, *160*(5), 865–875.

Voyich, J. M., Braughton, K. R., Sturdevant, D. E., Whitney, A. R., Said-Salim, B., Porcella, S. F., … DeLeo, F. R. (2005). Insights into mechanisms used by *Staphylococcus aureus* to avoid destruction by human neutrophils. *Journal of Immunology, 175*(6), 3907–3919.

Vuong, C., Saenz, H. L., Gotz, F., & Otto, M. (2000). Impact of the agr quorum-sensing system on adherence to polystyrene in *Staphylococcus aureus*. *The Journal Infectious Diseases, 182(6),* 1688–1693.

Walther, B., Monecke, S., Ruscher, C., Friedrich, A. W., Ehricht, R., Slickers, P., … Becker, Al. (2008). Comparative molecular analysis substantiates zoonotic potential of equine methicillin-resistant *Staphylococcus aureus. Journal of Clinical Microbiology*, *47*(3), 704–710.

Wang, R., Braughton, K. R., Kretschmer, D., Bach, T. H., Queck, S. Y., Li, M., … Otto, M. (2007). Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. *Nature Medicine, 13*(2), 1510–1514.

Weigel, L. M., Clewell, D. B., Gill. S. R., ClarK, N. C., Flannagan, S. E., Kolonay, J. F., … Tenover, F. C. (2003). Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science, 302*(5650), 1569–1571.

Xie, Y., He, Y., Gehring, A., Hu, Y., Li, Q., Tu, S. I., & Shi, X. (2011). Genotypes and toxin gene profiles of *Staphylococcus aureus* clinical isolates from China. *PLOS One, 6*(12), e28276.

Young, B. C., Golubchik, T., Batty, E. M., Fung, R., Larner-Svensson, H., Votintseva, A. A., … Wilson, D. J. (2012). Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. *Proceedings of the National Academy of Sciences in United States America, 109*(12), 4550–4555.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S. Lund, O., … Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *Antimicrobial Agents and Chemotherapy, 67*, 2640-2644.

Zhou, Y., Liang, Y., Lynch, H. K., Dennis, J. J. & Wishart D. S. (2011). PHAST: A fast phage search tool. *Nucleic Acids Research*, *39*(suppl 2), W347-W352.

ANNEX I

Table 8: Comparative table showing the virulence factors profile of each isolate along with the CC80 Reference Strain.

CC80-samples	**MB46**	**MB6**	**MB7**	**MB9**	**MB12**	**CC80 reference Strain**
Protein function	**virukence factors**	**virulence factors**	**virulence factors**	**virulence factors**	**virulence factors**	**virulence factors**
Intracellular adhesion proteins	*IcaA,IcaB, IcaC IcaD,IcapR*	*IcaA,IcaB,IcaC IcaD,IcapR*	*IcaA,IcaB, IcaC IcaD,IcapR*	*IcaA,IcaB,IcaC IcaD,IcapR*	*IcaA,IcaB,IcaC IcaD,IcapR*	*IcaA,IcaB,IcaC IcaD,IcapR*
extracellular adherence proteins	*eap*	*eap*	*eap*	*eap*	*eap*	*eap*
Ser-Asp rich fibrinogen-binding portein H	*sdrH, sdrD, sdrE, sdrC*	*sdrH, sdrE, sdrC*	*sdrH, sdrD, sdrE, sdrC*	*sdrH, sdrD, sdrE, sdrC*	*sdrH, sdrD, sdrC*	*sdrH, sdrD, sdrE, sdrC*
capsular polysaccharide biosynthesis protein	*cap5P,N, M,F,E,A/capA,capB*	*cap5P,N,M, F,E,A/ capA,capB*	cap5M,N, A,E,F,P/capA,B	*cap5P,N,M ,F,E,A/ capA,capB*	*cap5P,N,M, F,E,A/ capA,capB*	*cap5P,N,M, F,E,A/ capA,capB*
capsular polysaccharide synthesis enzyme	*cap8O,K, L,J,I,H,G*	*cap8O,K,L, J,I,H,G*	*cap8O,K, L,J,I,H,*	*cap8O,K,L, J,I,H,*	*cap8O,K,L, J,I,H,*	*cap8O,K,L, J,I,H,*
superantigen like protein	*set3,5,15/sal*	*set3,5,15/sal*	*set3,5,15/sal*	*set3,5,15/sal*	*set3,5,15/sal*	*set3,5,15/sal*
spa immunoglobulin G binding protein A	+	+	*?*	+	+	+
Thermonuclease	*nuc*	*nuc*	nuc	*nuc*	*nuc*	*nuc*
polysaccharide biosynthesis protein	*capD*	*capD*	*capD*	*capD*	*capD*	*capD*
type VII secretion protein	*esaA and essC*	*esaA and essC*	*esaA and essC*	*esaA and essC*	*esaA and essC*	*esaA and essC*
protein secretion	*essA*	*essA*	*essA*	*essA*	*essA*	*essA*

system						
ESAT-6/WXG100 family secreted protein EsxA/YukE	*esxA*	*esxA*	*esxA*	*esxA*	*esxA*	*esxA*
virulence factor EsxB family protein	*esxB*	*esxB*	*esxB*	*esxB*	*esxB*	*esxB*
EsaC protein within ESAT-6 gene cluster	*esaC*	*esaC*	*esaC*	*esaC*	*esaC*	*esaC*
Putative secretion accessory protein EsaB/YukD	*esaB*	*esaB*	*esaB*	*esaB*	*esaB*	*esaB*
putative secretion system component EssB	*essB*	essB	essB	essB	essB	essB
serine protease	*splF,A,B, C*	*splF,A,B,C*	splF,A,B, C	splF,A,B,C	splF,A,B,C	splF,A,B,C
lukF-PV, lukS-PV	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*	*LukF-PV, lukS-PV*
serine V8 protease	*sspA*	*sspA*	*sspA*	*sspA*	*sspA*	*sspA*
cysteine protease	*sspB,C,B2*	*sspB,C,B2*	*sspB,C,B2*	*sspB,C,B2*	*sspB,C,B2*	*sspB,C,B2*
hemolysin-delta/beta	*hld, hlb*	*hld, hlb*	*hld, hlb*	*hld, hlb*	*hld, hlb*	*hld, hlb*
alpha-hemolysin precursor	*hla*	*hla*	*hla*	*hla*	*hla*	*hla*
gamma-hemolysin component B precursor	*hlgB*	*hlgB*	*hlgB*	*hlgB*	*hlgB*	*hlgB*
gamma-hemolysin component C	*hlgC*	*hlgC*	*hlgC*	*hlgC*	*hlgC*	*hlgC*
gamma-hemolysin chain II precursor	*hlgA*	*hlgA*	*hlgA*	*hlgA*	*hlgA*	*hlgA*
clumping factor ClfB, fibrinogen binding protein	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*	*clfB,fib, clfA*

bifunctional autolysin Atl	*atl*	*atl*	*atl*	*atl*	*atl*	*atl*
von Willebrand factor-binding protein	*vwb*	*vwb*	*vwb*	*vwb*	*vwb*	*vwb*
cell wall associated fibronectin-binding protein	*ebh*		*ebh*	*ebh*	*ebh*	*ebh*
cell surface elastin binding protein	*ebpS*	*ebpS*	*ebpS*	*ebpS*	*ebpS*	*ebpS*
putative enterotoxin type A	*SEntA*	*SEntA*	*SEntA*	*SEntA*	*SEntA*	*SEntA*
exotoxin	*SExo*	*SExo*	*SExo*	*SExo*	*SExo*	*SExo*
exfoliative toxin A	*eta*	*eta*	*eta*	*eta*	*eta*	*eta*
leukotoxin LukE/ D	*lukE, LukD*	*lukE, LukD*	*lukE, LukD*	*lukE, LukD*	*lukE, LukD*	*lukE, LukD*
enterotoxin-like toxin	*SEnt-like*	*SEnt-like*	*SEnt-like*	*SEnt-like*	*SEnt-like*	*SEnt-like*
glycerol ester hydrolase	*geh*	*geh*	*geh*	*geh*	*geh*	*geh*
staphylocoagulase precursor	*coa*	*coa*	*coa*	*coa*	*coa*	*coa*
staphylokinase	*sak*	*sak*	*sak*	*sak*	*sak*	*sak*
complement inhibitor SCIN	*scn*	*scn*	*scn*	*scn*	*scn*	*scn*
triacylglycerol lipase	*lip*	*lip*	*lip*	*lip*	*lip*	*lip*
IgG-binding protein SBI	*isb*	*isb*	*isb*	*isb*	*isb*	*isb*
capsular polysaccharide synthesis protein	*capC*	*capC*	*capC,G*	*capC,G*	*capC,G*	*capC,G*
	emp	*emp*	*emp*	*emp*	*emp*	*emp*
Fibronectin binding	*FnbA and*	*FnbA and*	*FnbA and*	*FnbA and*	*FnbA and*	*FnbA and*

protein	*FnbB*	*FnbB*	*FnbB*	*FnbB*	*FnbB*	*FnbB*
fibrinogen-binding protein	*efb*	*efb*	*efb*	*efb*	*efb*	*efb*
Elastin binding protein	*ebps*	*ebps*	*ebps*	*ebps*	*ebps*	*ebps*

Table 9 Comparative table showing the plasmid content of each isolate along with its position, % identity and accession number. *High-scoring segment pairs.

	Plasmid	**%Identity**	**Query/ HSP * length**	**Contig**	**Position in contig**	**Note**	**Accession number**
MB46	rep10	100	477 / 477	scaffold_18	1008..1484	repL(pE5)	M17990
	rep21	95.27	803 / 999	scaffold_20	1580..2382	rep(pWBG754)	GQ900396.1
	rep7	99.88	821 / 821	scaffold_0	347681..348501	rep(MSSA476)	BX571857.1
	rep20	99.89	936 / 936	scaffold_12	8793..9728	repA(pWBG753)	GQ900395.1
MB6	rep7	99.88	821 / 821	scaffold_1	337273..338093	rep(MSSA476)	BX571857.1
	rep20	99.89	936 / 936	scaffold_16	5227..6162	repA(pWBG753)	GQ900395.1
MB7	rep7	100	496 / 570	scaffold_58	17..512	ORF(pKH1)	SAU38656
	rep7	99.88	821 / 821	scaffold_1	71556..72376	rep(MSSA476)	BX571857.1
MB9	rep7	99.88	821 /	scaffold	347748..34	rep(MSSA4	BX5718

			821	_0	8568	76)	57.1
	rep20	99.89	936 / 936	scaffold _16	5455..6390	repA(pWBG 753)	GQ9003 95.1
	rep10	100	477 / 477	scaffold _24	671..1147	repL(pE5)	M17990
MB1 2	rep20	99.89	936 / 936	scaffold _20	5226..6161	repA(pWBG 753)	GQ9003 95.1
	rep7	99.88	821 / 821	scaffold _12	707..1527	rep(MSSA4 76)	BX5718 57.1

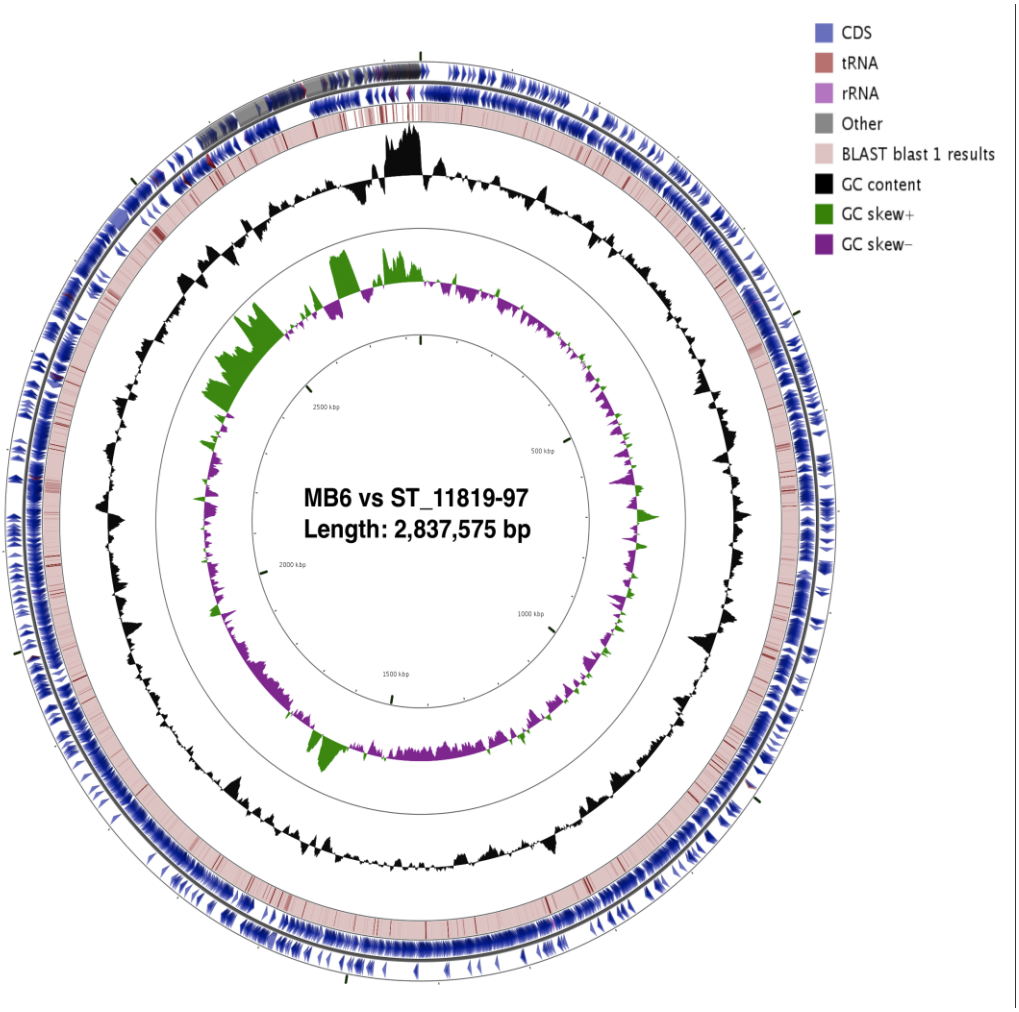

Figure 5 Genomic comparison between *S. aureus* MB6 and ST_11819-97. This graphical representation of the genome was generated using CGview server. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand, *S. aureus*

coding DNA sequence on the reverse strand, rRNA and tRNA are colored according to the legend inside the first two tracks, third track, which is the blast track (Reference strain 11819_97), represents the positions covered by the BLASTN alignment. Inside this track, white regions indicate parts of the input sequence that did not yield a blast hit, light pink represents parts of the input sequence that yield one blast hit, darker pink regions indicate parts of the input sequence that yield several blast hits (ie overlapping hits). These often include rRNA or tRNA genes or repetitive sequences for which several similar sequences in the CC80 reference strain are found.Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively.

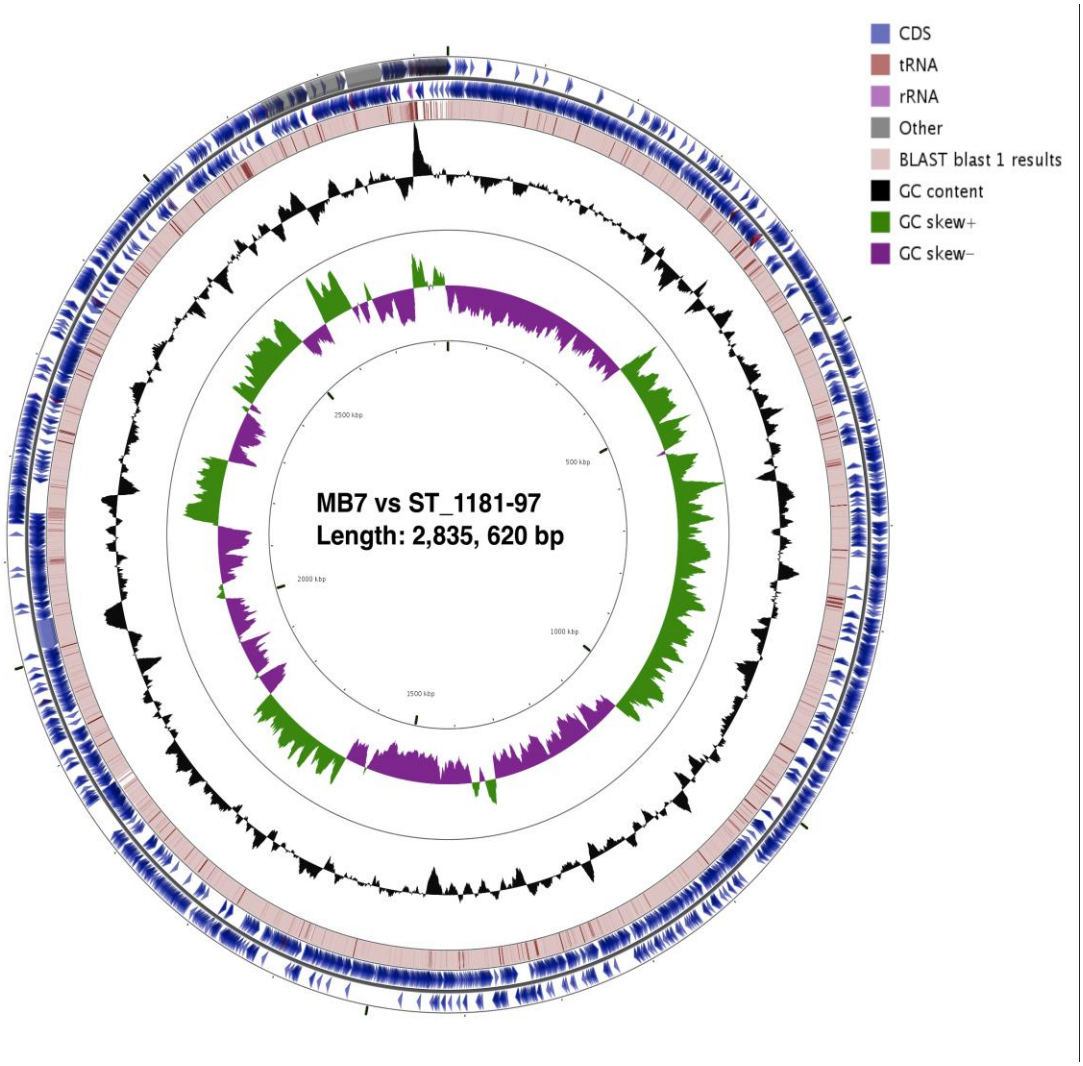

Figure 6 Genomic comparison between *S. aureus* MB7 and ST_11819-97. This graphical representation of the genome was generated using CGview server. From

outside to inside: *S. aureus*, coding DNA sequence on the forward strand, *S. aureus* coding DNA sequence on the reverse strand, rRNA and tRNA are colored according to the legend inside the first two tracks, third track, which is the blast track (Reference strain 11819_97), represents the positions covered by the BLASTN alignment. Inside this track, white regions indicate parts of the input sequence that did not yield a blast hit, light pink represents parts of the input sequence that yield one blast hit, darker pink regions indicate parts of the input sequence that yield several blast hits (ie overlapping hits). These often include rRNA or tRNA genes or repetitive sequences for which several similar sequences in the CC80 reference strain are found.Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively.

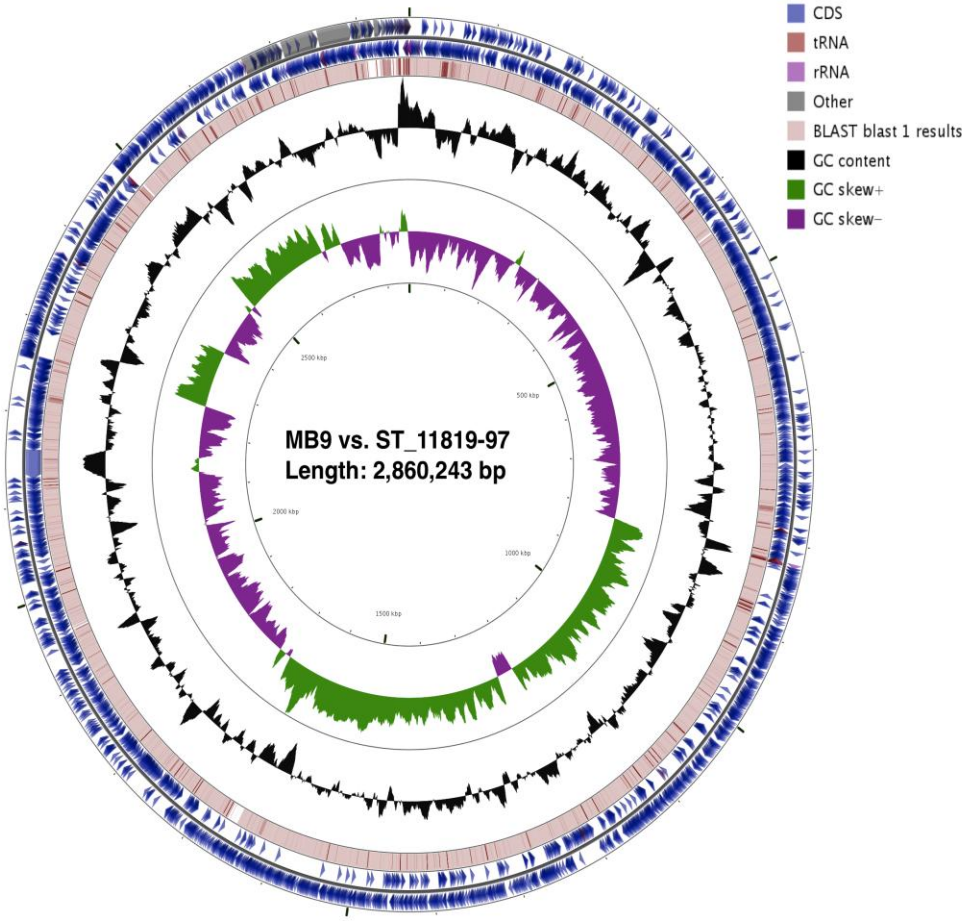

Figure 7 Genomic comparison between *S. aureus* MB9 and ST_11819-97. This graphical representation of the genome was generated using CGview server. From

outside to inside: *S. aureus*, coding DNA sequence on the forward strand, *S. aureus* coding DNA sequence on the reverse strand, rRNA and tRNA are colored according to the legend inside the first two tracks, third track, which is the blast track (Reference strain 11819_97), represents the positions covered by the BLASTN alignment. Inside this track, white regions indicate parts of the input sequence that did not yield a blast hit, light pink represents parts of the input sequence that yield one blast hit, darker pink regions indicate parts of the input sequence that yield several blast hits (ie overlapping hits). These often include rRNA or tRNA genes or repetitive sequences for which several similar sequences in the CC80 reference strain are found.Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively.

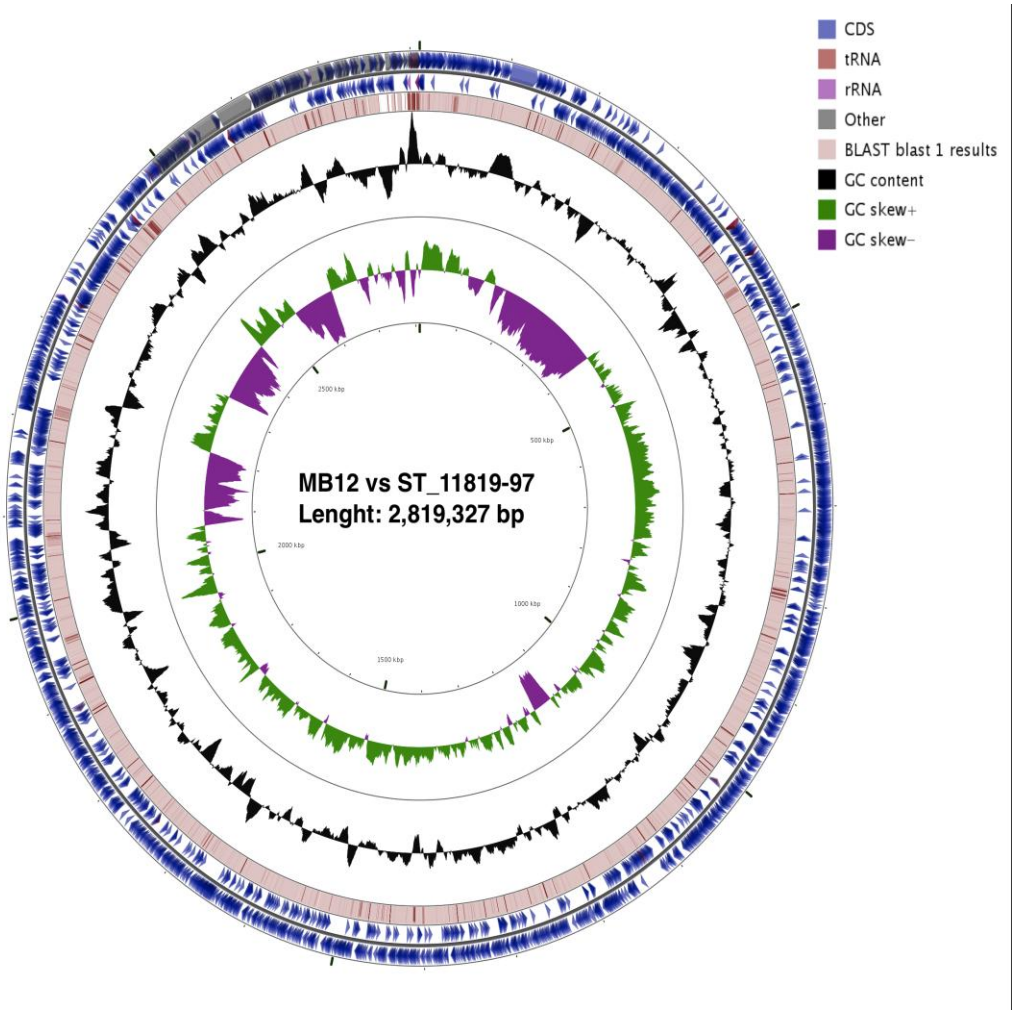

Figure 8 Genomic comparison between *S. aureus* MB12 and ST_11819-97. This graphical representation of the genome was generated using CGview server. From

outside to inside: *S. aureus*, coding DNA sequence on the forward strand, *S. aureus* coding DNA sequence on the reverse strand, rRNA and tRNA are colored according to the legend inside the first two tracks, third track, which is the blast track (Reference strain 11819_97), represents the positions covered by the BLASTN alignment. Inside this track, white regions indicate parts of the input sequence that did not yield a blast hit, light pink represents parts of the input sequence that yield one blast hit, darker pink regions indicate parts of the input sequence that yield several blast hits (ie overlapping hits). These often include rRNA or tRNA genes or repetitive sequences for which several similar sequences in the CC80 reference strain are found.Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively.

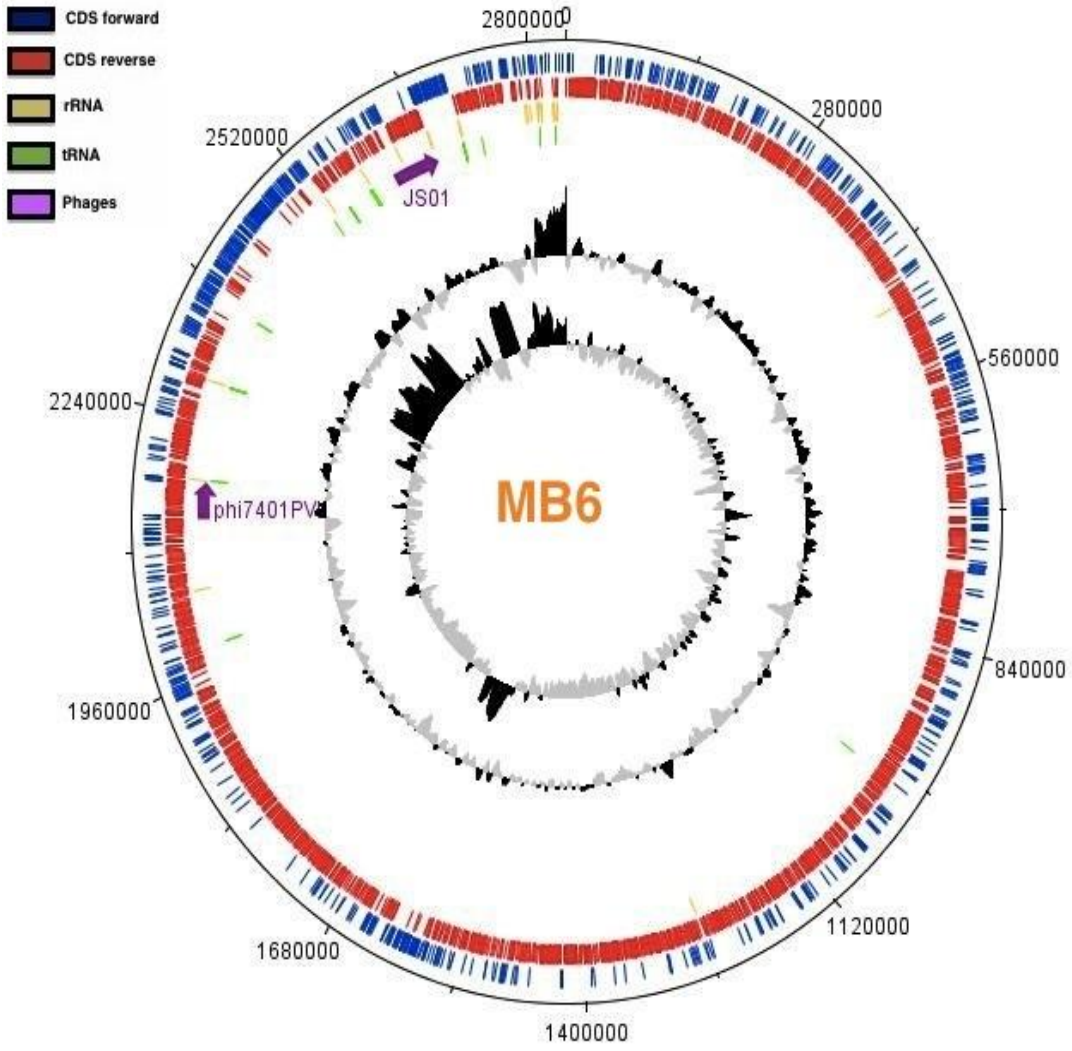

Figure 9 Genome atlas of *S. aureus* MB6. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand (blue); *S. aureus* coding DNA sequence2 on the reverse strand (red); ribosomal RNA genes (yellow); tRNA genes (green_ GC% (Black below mean and grey above mean); GC skew. The purple arrows represents the phages detected in the genome.

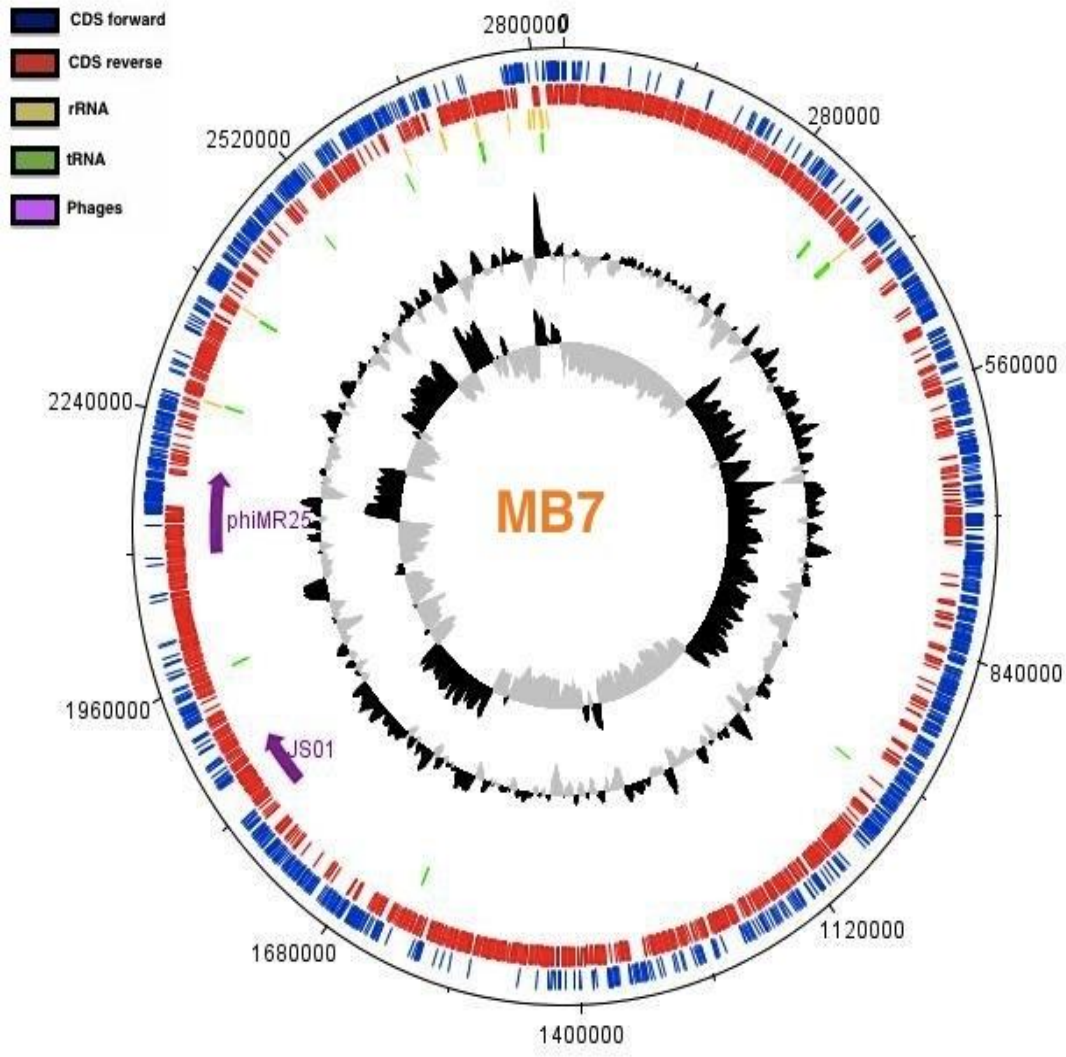

Figure 10 Genome atlas of *S. aureus* MB7. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand (blue); *S. aureus* coding DNA sequence on the reverse strand (red); ribosomal RNA genes (yellow); tRNA genes (green_ GC% (Black below mean and grey above mean); GC skew. The purple arrows represents the phages detected in the genome.

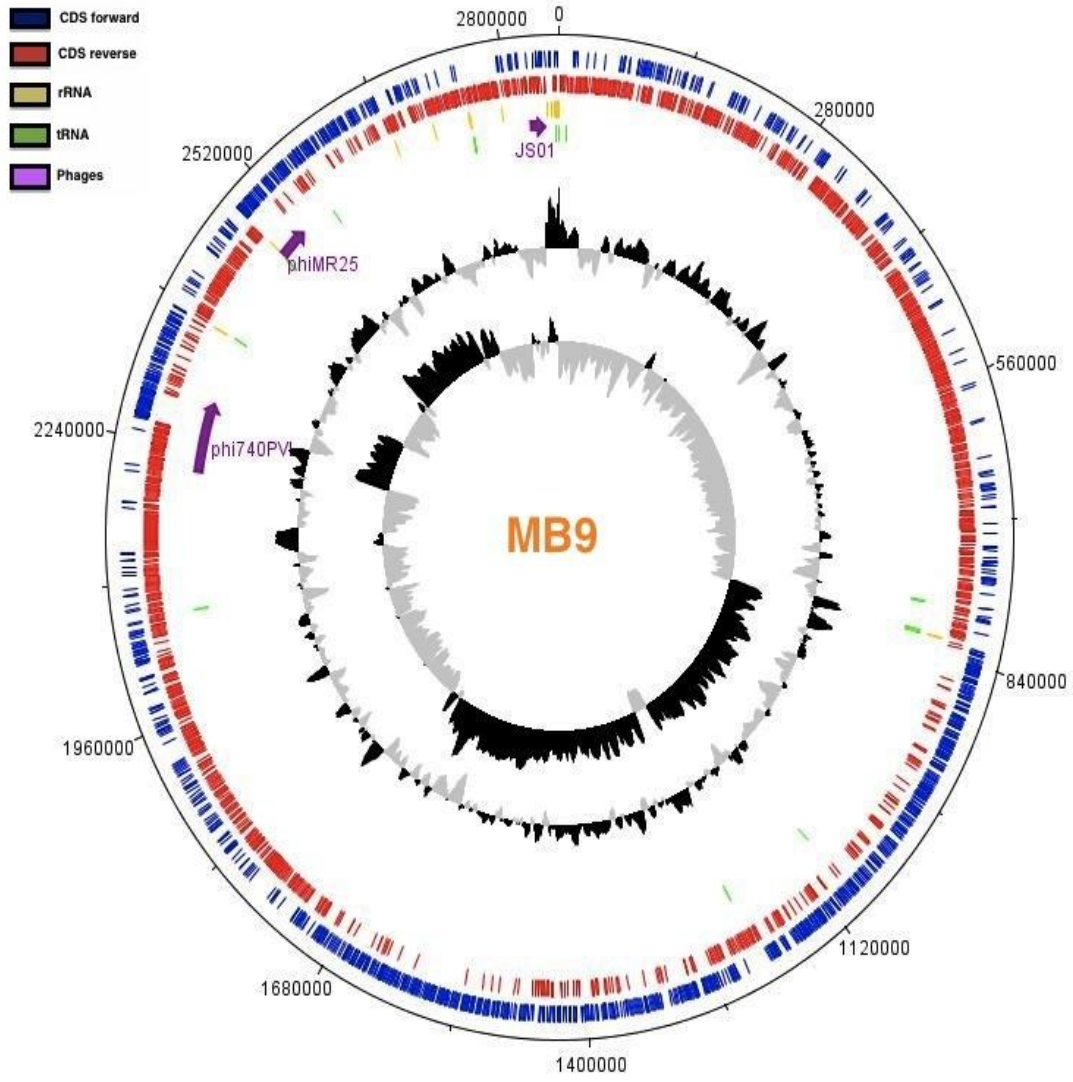

Figure 11 Genome atlas of *S. aureus* MB9. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand (blue); *S. aureus* coding DNA sequence on the reverse strand (red); ribosomal RNA genes (yellow); tRNA genes (green_ GC% (Black below mean and grey above mean); GC skew. The purple arrows represents the phages detected in the genome.

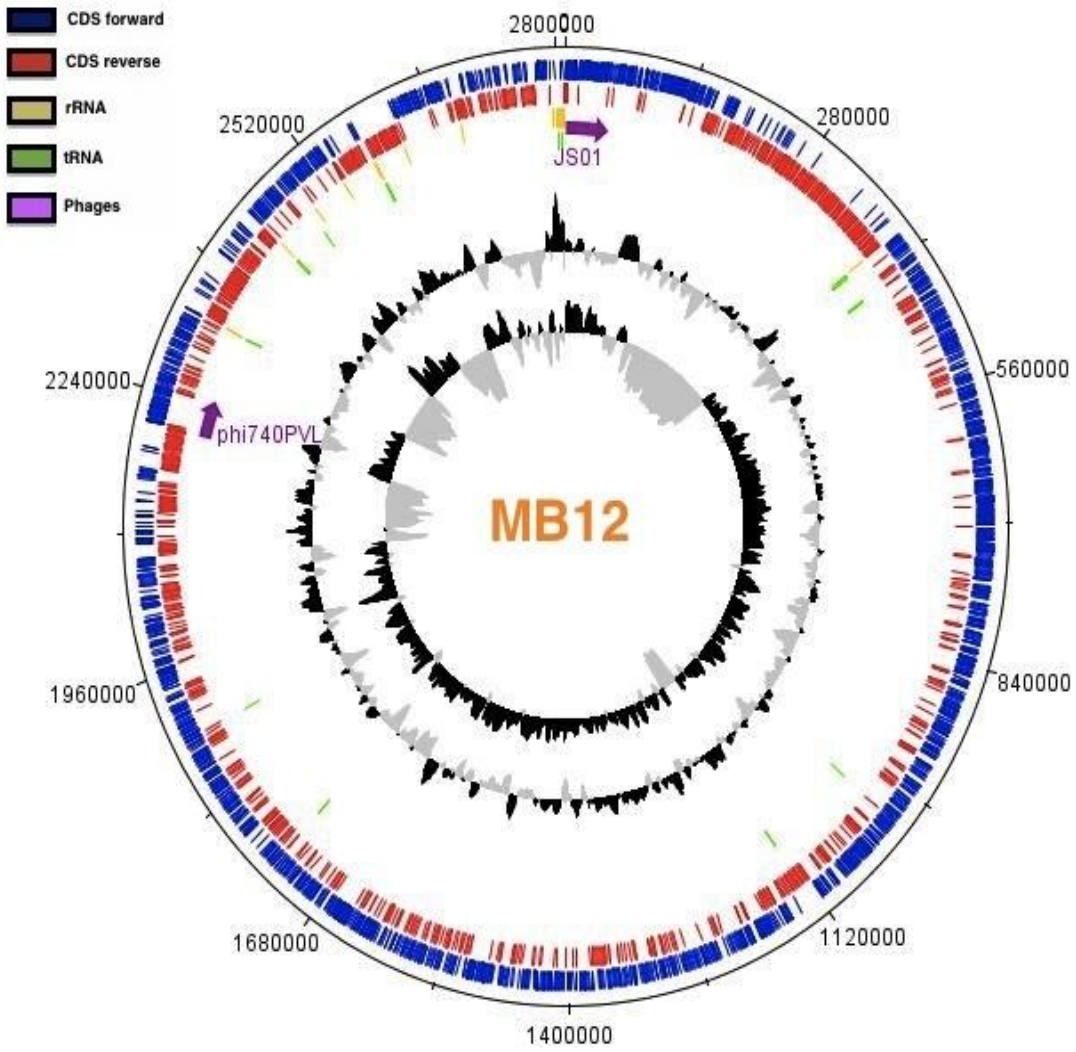

Figure 12 Genome atlas of *S. aureus* MB12. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *S. aureus*, coding DNA sequence on the forward strand (blue); *S. aureus* coding DNA sequence on the reverse strand (red); ribosomal RNA genes (yellow); tRNA genes (green_ GC% (Black below mean and grey above mean); GC skew. The purple arrows represents the phages detected in the genome.