12-1-2012

Cloning and Characterization of Hypothetical Exported Proteins from Community Associated Staphylococcus Aureus

Haninder Kaur
Indiana State University

VITA

Haninder Kaur

Education

Ph. D in Microbiology Jan, 2007 to Dec, 2012

Department of Biology, Indiana State University, Terre Haute, IN 47807, USA

M. Sc. in Microbiology July, 2003 to Jun, 2005

Osmania University, Secunderabad, Andhra Pradesh, 500003, India

B. Sc. in Microbiology, Botany, and Chemistry July, 2000 to May, 2003

Osmania University, Secunderabad, Andhra Pradesh, 500003, India

Publication

Kim, YK, Kaur, H, Ramalinga, AB, Dannelly, HK, Kahanov, L (2012) Staphylococcus aureus and MRSA in and around a therapeutic whirlpool in a college athletic training room, J Athletic Training, (in preparation).

Presentations (Selected)

Kaur, H., and McLean. M.A., *Cyanobacteria*, Unexpected microbes in acid mine drainage. *Poster presented at Research showcase, Indiana state university*. April 4, 2008.

Kaur, H., and Dr. Dannelly, H. K., Cloning and Characterization of SAS1678, a Hypothetical Exported Protein from Community-Associated *Staphylococcus aureus. Poster presented at Symposium celebration of student scholarship, Indiana State University.* April 21, 2010.

Kaur, H., and Dr. Dannelly, H. K., Cloning and characterization of hypothetical exported proteins from community associated strain of staphylococcus aureus. *Poster presented at IBASM: Indiana Brach of American society of Microbiology Annual Meeting.* April 21, 2012.

CLONING AND CHARACTERIZATION OF HYPOTHETICAL EXPORTED PROTEINS FROM COMMUNITY ASSOCIATED *STAPHYLOCOCCUS AUREUS*.

A dissertation

Presented to

The College of Graduate and Professional Studies

Department of Biology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Haninder Kaur

December 2012

Keywords: MSSA476, CA-MRSA, Keratinocytes, qPCR, Gateway cloning

COMMITTEE MEMBERS

Committee Chair: H. Kathleen Dannelly, PhD

Associate Professor, Biology: Microbiology

Indiana State University

Committee Member: Swapan K. Ghosh, PhD

Professor, Biology: Immunology

Indiana State University

Committee Member: Gary W. Stuart, PhD

Professor, Biology: Molecular and Developmental Genetics

Indiana State University

Committee Member: Allan Albig, PhD

Assistant Professor, Biological Sciences: Biomolecular Sciences

Boise State University

Committee Member: Margaret M. Moga, PhD

Associate Professor, Anatomy and Cell Biology

Indiana University School of Medicine

ABSTRACT

Methicillin resistant *Staphylococcus aureus* (MRSA), a major cause of nosocomial infections, has acquired resistance to beta-lactam and other antibiotics. Recently, community associated MRSA (CA-MRSA) has developed independent of hospital associated MRSA (HA-MRSA). One of the major differences between the hospital and community strains is that the former is multi-resistant to antibiotics while the latter is not as resistant but is significantly more invasive. This increased invasiveness and the ability to cause life-threatening infections, even in immunocompetent individuals, makes CA-MRSA critically important as a public health problem. CA-MRSA is known to cause skin and soft tissue infections; bacteria interact with host skin cells and gain access to deeper tissues causing invasive infections. During this process the bacteria may secrete proteins that aid in the interaction with the host by adhering, invading or causing host cell death and lysis.

To understand the virulence mechanisms involved in invasion, we investigated genes described as hypothetical proteins in MSSA476. The bioinformatics-selected proteins showed high probability of being secreted and most were unique to CA-MRSA. Our analysis showed 24 such genes. This study shows primer design for 15 of the genes (7 of the 24 had already been cloned in our laboratory). Using gateway cloning, the 15 genes were cloned into BL-21 expression clones.

CA-MRSA's are known for causing invasive skin infections. To further understand the involvement of our proteins of interest in invasion, human keratinocyte cell lines were used in a

study of virulence and interaction with skin. To understand the involvement of our hypothetical secreted proteins, we investigated the mRNA expression level, using RT-qPCR and Livak method, of 20 hypothetical exported proteins in presence of human dermal keratinocyte cell line. Our investigation revealed two genes that showed increased mRNA expression in the presence of keratinocytes, which may be due to factors associated with keratinocytes that may have triggered increased mRNA expression. Keratinocytes are capable of forming cell-cell junctions and producing antimicrobial peptides and cytokines in response to microbes. The increased mRNA expression of two genes may be towards binding to junctions for invasion or may be expressed in response to antimicrobial peptides or cytokines.

PREFACE

The aim of this study is to understand the involvement of hypothetical exported proteins, from community associated *Staphylococcus aureus*, in causing invasive infection that community associated methicillin resistant *S. aureus* (CA-MRSA) are known for. We investigated mRNA expression of hypothetical exported proteins in presence of human keratinocytes dermal cell lines. Our study shows increased mRNA expression of two genes in presence of keratinocytes, which may indicate possible interaction of these hypothetical exported proteins with keratinocytes.

ACKNOWLEDGMENTS

My most sincere thanks to my advisor Dr. H. Katheleen Dannelly for her guidance and support during the course of my PhD program. It is not often to find an advisor who is always available to listen and discuss all the little issues that come up in the course of research. My thanks to my committee members: Dr. Gary W. Stuart, Dr. Swapan K. Ghosh, Dr. Allan R. Albig and Dr. Margaret M. Moga for their insight and valuable suggestions with research.

My thanks to all members of Dr. Dannelly's lab: Deepak for his support, encouragement and teaching me lab techniques, Anupama for support and research related discussions, Toru for support and for all the fun times we had in the lab and Lauren for helping me with the gateway cloning.

My thanks to Dr. Michael W. King, Indiana University School of Medicine for allowing us to use qPCR equipment.

I thank Bikram Sharma for his help with understanding qPCR and for teaching me how to analyze qPCR results.

I thank my friends Reema, Neha, Gurinder, and Sreejith for their support and encouragement.

I would like to thank my family members, husband Arun Seetharam, daughter Avani Seetharam, father Bhupender Singh, mother Amritpal Kaur, brother Jagdeep Singh and sister's, Ritender Kaur and Jaspreet Kaur, for their understanding, love and encouragement during the course of my PhD program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Staphylococcus Aureus

Staphylococcus aureus is gram positive cocci in clusters. It is a commensal found in 20% of the population as carrier and 60% of the population as transient carrier on the skin and nares [1, 2]. It is normal flora of the skin and mucous membrane. It can colonize the host without causing any signs or symptoms of infection and, at the same time, with various mechanisms to defend the host immune system, *S. aureus* can become pathogenic when it gains access through cuts, abrasions or other openings [3]. It is one of the most common skin-infecting organisms in the United States. Skin-related infections include abscesses, cellulitis, folliculitis, carbuncles, furuncles, impetigo, and scalded skin syndrome [3, 4]. *Staphylococcus* infections are not limited to skin infections; it can also cause bacteremia, sepsis, pneumonia, meningitis, endocarditis, toxic shock syndrome, and osteomyelitis [3, 5, 6].

Staphylococcus aureus can easily be transmitted from an infected person to a healthy individual through direct skin-to-skin contact or sharing objects that can carry bacteria, for example towels, lip balm, etc. [7]. There is a high risk of infections in closed settings such as schools, daycares, and athletic training centers where human interactions are common [3].

Methicillin resistant *Staphylococcus aureus* (MRSA) has acquired a staphylococcal cassette chromosome mec (SCCmec) that gives resistance to a wide range of beta-lactam

antibiotics [9]. SCCmec is a mobile genetic element that can be acquired by horizontal gene transfer (HGT). Many types of SCCmec have been identified, and depending on the type of SCCmec, the strains are classified into two types: hospital associated (HA-MRSA) and community associated (CA-MRSA). Various strains have been identified based on the antibiotic resistance acquired by this bacterium, such as methicillin resistant *S. aureus* (MSRA) and methicillin susceptible *S. aureus* (MSSA), vancomycin resistant *S. aureus* (VRSA), vancomycin intermediate *S. aureus* (VISA) [10].

Hospital Associated Methicillin Resistant Staphylococcus Aureus (HA-MRSA)

HA MRSA causes nosocomial infections and is acquired while admitted to a hospital or healthcare facility [11]. HA-MRSA can easily be transmitted by infected instruments or catheters used during surgery that can deposit the bacteria to the surgery site and can cause systemic infection [7]. It can also be transmitted from one patient to another patient or patient's visitors may acquire it [12]. In some cases, the patient is a carrier of MRSA [1, 13]. Soldiers can carry it along with mud and debris collected in the clothes, and during surgery these bacteria get transmitted to surgical wounds [1]. In many instances, MRSA infection is not detected while the patient in admitted in the hospital; signs and symptoms appear immediately, or a few months, after hospitalization [14].

Symptoms and infections associated with HA MRSA are identical to *S. aureus* infections; the main threat with HA MRSA is the acquisition of antibiotic resistance genes which makes it difficult to treat [3]. HA MRSA has acquired SCCmec I, II and III [9] giving it resistance towards a wide range of antibiotics. SCCmec II and III are larger in size and are rarely transferred by HGT. Transmission of antibiotic resistance is mostly due to selective pressure and

vertical gene transfer [15]. With continuous use and exploitation of antibiotics, more resistant strains like VISA and VRSA have been identified in recent years [16].

Vancomycin Resistant Staphylococcus Aureus (VRSA)

Vancomycin is a glycopeptide antibiotic used to treat patients with MRSA infections. The *van*A gene codes for vancomycin resistance and is located on a transposon that has been acquired by HGT from *Enterococcus faecalis* [17]. In the case of VISA, a thicker cell wall does not allow vancomycin to get further than outmost cell wall layers [18].

Community Associated Methicillin Resistant Staphylococcus Aureus (CA-MRSA)

CA MRSA mostly causes skin and soft tissue related infections [16]. It occurs commonly in athletes, especially contact sports, and in "closed" societies such as prisons and military personnel [3, 19, 20]. However, CA MRSA also occurs sporadically in the general population, and this occurrence is on the rise [16].

Most *Staphylococcus* infections are minor skin infections and may even be self-limiting [16]. But in the case of CA MRSA, the infectious agent is highly invasive [16]. CA MRSA causes skin infections when it comes in contact with open wounds, cuts and abrasions, and can be transferred very efficiently via skin to skin contact and indirectly via fomites [3]. Infections often start as a red pimple that may appear as a spider bite, or more seriously, as a boil [16]. This apparent benign beginning is followed by fever and rash; the lesion grows bigger and more painful, and eventually opens to reveal the pus-filled lesion [16]. The infection should be treated by draining the lesions and use of appropriate antibiotics [3]. Increased invasiveness and the ability to cause life threatening infections, even in immunocompetent individuals, distinguishes CA MRSA as a critically important public health problem [21]

CA MRSA acquired genes and spread in the community independent of HA MRSA; genetically they are distinct [22]. One major genetic difference is that CA MRSA has acquired a different set of SCCmec genes from HA MRSA and does not possess a wide range of antibiotic resistance [16]. CA MRSA possesses SCCmec IV and V [9, 10] giving it resistance to some of the beta lactam antibiotics; these SCCmec genes can be easily transferred by HGT due to their small size [10].

Methicillin Susceptible Staphylococcus Aureus Strain 476 (MSSA476)

MSSA476 is a methicillin susceptible stain reported to have caused infection in an immunocompetent child [23]. It has acquired a different SCCmec, $SCCmec_{476,}$ that gives the organism resistance to fusidic acid. It was the acquisition of SCCmec by MSSA476 from other species of *Staphylococcus* that lead to the spread of the SCCmec gene among CA-MRSA. The SCCmec acquired by MSSA476 is similar to those from *Staphylococcus epidermidis* and *Staphylococcus hominis* , suggesting interspecies HGT [24, 25].

MSSA476 is highly similar and non-distinguishable from the CA MRSA MW2 strain; multilocus sequence typing (MLST) revealed that both strains belong to ST1 type [23]. They do, however, have dissimilarities; SCCmec IV and the PVL genes are absent in MSSA476 [23]. Due to the high similarity between MW2, a community associated strain, and MSSA476, studying hypothetical exported proteins from MSSA476 that are 100% identical to those found in MW2 would aid our understanding of the CA MRSA strains. The invasive nature of CA MRSA may be associated with one or more of these hypothetical exported proteins. Previous study in our laboratory has revealed a hypothetical exported protein from MSSA476 inhibited *Caenorhabditis elegans* growth and showed signs of toxicity [26], stressing the importance of studying

hypothetical exported proteins from MSSA476 to understand the mechanisms by which these organisms are able to cause infection and deep tissue damage.

Skin as Barrier - Keratinocytes

Skin is the outer layer covering our body and accounts for 16% of body weight; it is the largest organ of human body [27]. The main function of skin is to keep organs and tissues intact in the body and to keep them safe from foreign objects. Skin is made of many layers of cells that are broadly classified as epidermis and dermis [27]. The epidermis is the outermost layer of skin and is comprised of 4 - 5 layers depending whether it is a thick layer of epidermis, such as on the palm, or thin layer of epidermis, such as the eyelids. The layers are as follows: stratum basale, stratum spinosum, stratum granulosum, stratum lucidum and stratum corneum [28].

The epidermis plays a role in retaining water, blocking toxins and harmful chemicals, and keeping microbes out of the body. The majority of epidermal layers are comprised of keratinocytes from stratum basale to stratum granulosum [28]. At the stratum corneum, the keratinocytes become corneocytes [29].

Stratum basale is the innermost layer of epidermis. Cells in stratum basale are dividing and non-dividing cuboidal keratinocytes and are responsible for renewing epidermal cells. As the keratinocytes divide by mitosis, the new cells take the place of the old cells, pushing the old cells outward [29]. It takes about 28 days for the cells to divide and move outward from stratum basale to stratum corneum [31]. As the cells are pushed from the basale layer to the stratum spinosum, they attain a polygonal shape and start synthesizing keratin. The next layer of cells, stratum granulosum, has lost its nuclei and organelles [31]. The main function of these cells is to produce keratin and lipids. The stratum corneum is the outermost layer of the epidermis and consists of dead cells that are packed with keratin and lipids [28].

Lipid lamella consists of free fatty acids that bring the pH of the skin down and inhibits bacterial colonization [32]. The stratum corneum acts as a barrier and retains water. The cells of the stratum corneum are replaced by desquamation of these cells; the desmosomes are dissolved enzymatically, breaking bonds between two cells to release them to the environment, and thus relocating bound microbes [28, 29].

Skin comes into contact with various microbes daily, from benign normal flora to pathogens like *Staphylococcus sp, Streptococcus sp.,* causing little or no harm to healthy individuals [34]. Everyday activity can cause damage to the outermost layers of the epidermis and open passages for microbes to gain access to inner cells and tissues. Usually, keratinocytes are the first line of cells a bacterium must overcome, depending on the size of the break in the skin. Keratinocytes are capable of detecting bacteria via toll-like receptors [35] and produce antimicrobial peptides giving protection from some bacteria. Keratinocytes are also capable of producing cytokines at low levels at all times[32]. However, in the case of injury, toxins, or other foreign materials, keratinocytes produce elevated levels of cytokines that attract macrophages and neutrophils to the site of breech [36, 37].

Tight Junction

Tight Junctions (TJ) play an important role in the skin as a barrier towards chemicals and microbes and in retaining water. Loss of the outermost cells, the stratum corneum, does not cause as much harm as the loss of layers with TJs along with keratinocytes or the basale layers [32]. The epidermis is held together with TJs and desmosomes. TJs are formed by proteins bound to each other and connected to cells by actin fibers. TJ associated proteins are found throughout the epidermal layer but are mostly found in the stratum granulosum layer [32]. The TJs are capable

of holding cells close together and not allowing bacteria to pass. If a pathogen does have to pass through, it must invade the epithelial cells or find its way in through a cut or abrasion [28].

TJs in the epithelial cells allows passage of small molecules or ions through the paracellular pathway [39]. Claudins and occludins contain N and C terminal with 4 transmembrane domains that forms two loops in the paracellular region. The loops from two cells interact with each other and form pores that allow small molecules to pass through [39, 40]. Claudins are found in epidermal layers, hair follicles, and endothelial cells [32], whereas occludins are specific to tight junctions [40]. Claudin isotypes play roles in tightening and pore forming. Tightening claudins isoforms are 1, 4, 5, 8, 11, 14, and 19 and pore forming are 2, 7, 10, & 15 [39]. Claudin 1 and 4, tightening claudins, are found in the stratum granulosum. Claudin 7 is found in the basal layer and other layers of epidermis. It is speculated that claudin found in cells without TJ plays role in proliferation [39]. Mice, deficient in claudin 1 or desmosome 3, would not survive. Both of these cell-cell adherence junctions are important in retaining water and maintaining healthy stratum corneum [32]. A study showed that death of neonatal mice with claudin 1 deficiency was due to transepidermal dehydration [39].

Adherence Junctions

Both adherence junctions and desmosomes are involved in binding cells together using the cadherin family of proteins (calcium dependent) [38]. Adherence junctions are structures that form networks with the actin cytoskeleton lending structural integrity [41]. They are not only involved in cell-cell interactions and barrier function but also play a role in regulating other processes associated with epidermal cells such as shape, growth and division [41]. P-cadherin is expressed in hair follicles and basal layers, and E-cadherin is expressed throughout the epidermis. Young et al. Showed that loss of E-cadherin in keratinocytes affected adherence

junctions and differentiation of keratinocytes in the epithelia [42]. E-cadherin forms transmembrane junctions with α and β-catenin anchoring protein. Cadherin dimers form bonds with anchor proteins that in turn bind to the intercellular actin cytoskeleton, forming intercellular adherence junctions [38, 42].

Desmosomes

Desmosomal cadherins stabilize adhesion between cells in the epidermis [41]. Similar to adherence junctions, desmosomal cadherins form transmembrane junctions between epidermal cells. Desmocollins and desmogleins are cadherins involved in forming intercellular heterodimers in adjacent cells [32]. The cytoskeleton is attached to the cadherin heterodimers by plankoglobin and desmoplakins, anchor proteins or desmosomal plaques [32, 38]. These junctions are highly resistant and flexible in the epidermal region where the keratinocytes are constantly differentiating and require desmosomes to withdraw to allow cells in the epidermis to differentiate, changing their size and shape [32]. In the stratum corneum, desmosomes form corneodesmosomes that are modified desmosomes, the plaque gets embedded into the plasma membrane and the cytoskeleton is mostly keratin bundles. The desquamation of the stratum corneum is due to proteoltic degradation of desmosomes [32].

Antimicrobial Peptides

Keratinocytes can produce cationic antimicrobial polypeptides as part of the innate host defense, capable of inhibiting a wide range of microorganisms. β-defensins, cathelicidin LL-37 and RNase7 are the families of antimicrobial peptides produced by keratinocytes [43, 44]. The mode of action is mainly creating pores on target cells or structures that makes the microbial membrane permeable [44], but more recently it was found that antimicrobial peptides can

translocate into the bacterial cytoplasmic membrane and are involved in inhibiting cell processes [45]. The target for these peptides are negatively charged microbial cells; human cells are also negatively charged, but the high amounts of cholesterol lipids in the epidermal layer prevents these antimicrobial peptides from forming pores in eukaryotic cells [46].

Defensins

α-defensins are primarily found in the neutrophil granules and paneth cells where they play a role in the non-oxidative pathway to kill bacteria and thus control bacterial number [32] whereas β-defensins are expressed in epithelial cells and circulating cells [32].

Calcium plays an important role in the differentiation of keratinocytes and the regulation of β-defensins. In vitro studies with keratinocytes required increased amount of calcium chloride in the media to initiate keratinocyte differentiation [43]; the same was seen in the epithelial cells of normal skin with an increased calcium gradient.

There are 4 types of β-defensins (hBD-1, hBD-2, hBD-3and hBD-4) [43, 44]. hBD-1 is active against Gram negative bacteria and is constitutively produced in small amounts, therefore, inhibiting Gram negative bacteria from colonizing the skin [32, 44]. hBD 2, 3, and 4 are induced by inflammation [32]. hBD-2 is localized in the upper epithelial stratum corneum layers and shows strong antimicrobial activity towards Gram negative bacteria and low activity towards Gram positive bacteria [36, 42]. Increased calcium efflux at the site of inflammation and sites where the skin barrier is disrupted would enable hBD-2 to express and be delivered to the site for added protection against invading microbes [32]. hBD-3 shows broad spectrum antimicrobial activity and is induced by tumor necrosis factor α (TNF-α), interleukin-1 (IL-1) and IL-6 in response to Gram positive and Gram positive bacteria [44]. hBD-4 shows broad spectrum

antimicrobial activity at micromolar concentrations and is induced by PAMPs and cell differentiation [44].

Cathelicidin

Cathelicidin are expressed in epithelial cells, mast cells and neutrophils [32]. Cathelicidin LL-37 is the type found in humans and is not only responsible for the antimicrobial activity towards a broad spectrum of microbes but also works in synergy with other antimicrobial peptides such as β-Defensins against invading microbes [44]. Studies have shown antimicrobial activity against *S. aureus* with LL-37 in association with hBD-3 and lysozyme [47]. Cathelicidins are constitutionally expressed in sweat glands, hair and nails, fetal and neonatal skin, giving protection from bacteria to the fetus while the adaptive immune system is still developing [32]. Cathelicidins usually have a 30 amino acid signal peptide, highly conserved 100 amino acids sequence, and a C terminal peptide that varies in size. The active peptide is cleaved by proteolysis, binding to the bacterial membrane, disrupting it or inhibiting protein synthesis prior to disruption [32].

RNase7

Another type of antimicrobial expressed by keratinocytes is RNase 7. High expression of RNase 7 is observed in the stratum corneum [44, 48]. It has ribonuclease activity and has been found effective against a broad spectrum of microbes [48]. Studies showed that incubation of RNase 7 with pathogens for 3 hrs resulted in at least a 2-log reduction of *Enterococcus faecium*, *Pseudomonas aeruginosa*, *Propionibacterium acnes, E. coli*, MRSA and *Candida albicans* [49]. In vivo experiments showed that the amount of RNase secreted by skin at different sites on the body varied between 3-12.5 ng/ml [46]. While studies show high expression of RNase 7 in primary keratinocytes compared to other antimicrobial peptides [48], there are no apparent toxic effects on the keratinocytes [49]. Immunohistochemistry performed with RNase 7 antibody in

healthy skin revealed the presence of RNase 7 throughout epidermal layers, with an increased gradient from the basal layer to the outermost stratum corneum layer [48]. Studies have shown mRNA expression of RNase 7 in the skin, respiratory and urinary tract and low expression in the gastrointestinal tract [49].

Cytokines

In keratinocyte culture, interleukin-1 (IL-1), IL-6 and tumor necrosis factor α (TNF-α) are produced in response to bacterial invasion or injury [32, 44]. These proinflammatory cytokines are also involved in activating antimicrobial peptides like defensins and LL-37 [44]. There are two types of IL-1: IL-1α and IL-1β. IL-1α is expressed by many cells but in keratinocytes IL-1α is constitutively expressed at high levels [32]. Both differentiated and undifferentiated keratinocytes are capable of producing IL-1, but cultured keratinocytes produce IL-1ra, a form that is not secreted and remains inside the cells due to additional amino acids at N terminus and the lack of a signal sequence [32]. IL-1α and IL- 1β are present in epidermal layers but are present in high concentrations in the outer epidermal layers [32]. Preformed IL-1 in the epidermis can initiate an immune response in case of injury. TNF-α is induced by endotoxins and UV light and is produced in culture, which can aid in the differentiation of keratinocytes [32]. The primary source of IL-6 is epidermal keratinocytes. In normal skin cells, IL-6 is located in the basale, granular and horny layers but not in the stratum spinosum layer [32].

Bacterial Proteins

Identifying virulent genes requires identifying genes specific to a pathogen, meaning it should be limited to the pathogenic strain to be investigated and not shared with non-pathogenic strain [50]. Even though some bacterial strains harbor similar genes, the regulation of these genes

plays a vital role; for example, *Salmonella enterica* and *Shigella flexneri* possess similar invasion genes that are homologous, but expression is regulated by oxygen in one [51] and temperature in the other bacteria [52]. Therefore,, genes may be expressed at varying levels and conditions, differently by both or more pathogens sharing similar genes.

Pathogens are capable of producing a wide range of proteins, each with its own destination [50]. This destination can be the cytoplasm, bacterial surface, secretion outside the bacterial cell, or secretion directly into a eukaryotic cell. Whatever the destination, the protein needs to be accurately directed to function correctly. Proteins secreted outside the cell must have signal peptides to direct them; with Gram positive cells, the protein must to pass through only one lipid bilayer, whereas in Gram negatives, it must has to pass through both the cytoplasmic and the outer membranes to exit the cell, making it a more complicated process.

When microbes invade a host, the defense system is triggered, and the immune response to the foreign object is activated. It is the normal reaction of a host to a foreign object, and the microbes have developed many tactics for evasion [54]. To do so, many bacterial proteins are released into the host system that play an important role in breaching host defense mechanisms and aiding the pathogen in colonization, adherence, etc. For example, some microbes are capable of removing heme from hemoglobin to acquire iron for replication [55], and it is important for a pathogen to not only survive but also to propagate within the host. *Staphylococcus aureus* secrets many such proteins that are toxic and cause host cell damage [6].

The entire genome of *S. aureus* MSSA476 is available [23], and this has given the opportunity to study the strain in great detail. In the sequence analysis, many unknown or hypothetical genes with no known function were discovered. Further computational analysis has identified within those hypothetic genes, hypothetical exported proteins from MSSA476 that

may be enzymes or proteins that are virulent in nature, bind to the substrate, interact with host cells, or interact with the host immune system. Identifying and defining the hypothetical exported protein from MSSA476 will help us understand the MSSA476 and, thus, to better understand CA-MRSA and its extreme invasiveness.

Caenorhabditis Elegans

Caenorhabditis elegans is a eukaryote, very easy to grow and maintain in the laboratory, and has been used as a model organism since the introduction of *C. elegans* by Dr. Sydney Brenner in 1963 [56]. *Caenorhabditis elegans* is a non-infectious organism that can be grown in petri dishes with nematode growth media (NGM) and *E. coli* as food sources (Fig 1). The life cycle of *C. elegans* is 3-5 days and each adult is capable of producing 200 eggs. In a short period of time, high numbers of nematodes can be collected for use in laboratory tests.

Caenorhabditis elegans is attracted to certain chemicals including bacterial byproducts; these products act as cues used by the nematode to detect bacteria in the vicinity. Some chemicals act as repellents, such as bacterial toxins or other harmful chemical that are detected by the animal's neurons. *Caenorhabditis elegans* is also capable of showing variation in morphology and behavioral patterns based on the environment, temperature, touch, etc. [57]. Therefore, growing *C. elegans* in the presence of proteins that are toxic in nature, these worms show predicted behavioral patterns or morphologies that will help us understand the test protein.

Figure 1 NGM media plate with *C. elegans* feeding on *E. coli* lawn.

A single adult nematode is highlighted within a blue square.

CHAPTER 2

METHODS

Computational Analysis

Genes described as hypothetical proteins were identified using the Gene Database website (Gene DB) and genes with probability of being hypothetical exported proteins were identified using five software programs: Phobius [58], SignalP2 version [59], Signal P 3 version [60], SPOCTOPUS [61] and TMHMM2.0 version [62]. Each of these programs is capable of identifying signal peptides, transmembrane domains, and the cytoplasmic or non-cytoplasmic nature of proteins, In addition Signal P and Phobius are capable of identifying signal cleavage site.

Proteins identified as hypothetical exported proteins (Table 1) were further analyzed using BlastP, which is capable of aligning protein sequences with a protein database; therefore, it determines the percent identity of the protein of interest as compared to proteins found in the database. Percent identity or homology of one protein with another can help in determining structural and functional domains in the test proteins. Percent identity would also help in determining if the proteins are unique to CA-MRSA's.

Primers Design

Primers were designed for 14 genes using a primer design tool from Invitrogen (Table 1). Twenty nucleotides upstream of the start codon were included when designing primers including a ribosomal binding site (RBS) (coded AGGAGG). Including the RBS ensures that the protein is expressed after acquiring expression clones. Excluding the RBS would reduce the chances of protein expression as the RBS on the vector may fall too far from the start codon. Both forward and reverse primers had additional 25 bp attachment sites (attB1 and attB2 sites respectively) [63]. The attB sites aided in the BP and LR recombination reaction's, in presence of clonase II enzymes (the principle behind the Gateway cloning technique).

When designing primers, the stop codon is excluded to ensure that the V5 epitope and His tag from the expression clone are fused to the protein. Fusion proteins can be purified using nickel column chromatography for His-tag fusion proteins or anti-V5 epitope antibody can be used for affinity column chromatography.

On the reverse primer, a 7 amino acid sequence coding the AcTEV protease site (Glu-Asn-Leu-Tyr-Phe-Gln-Gly) was added between the attB site and reverse primer region complementary to the gene (Table 1). The AcTEV site is specifically recognized by AcTEV protease, which cleaves it between 6^{th} (Gln) and 7^{th} (Gly) amino acid. After protein purification and dialysis of protein, AcTEV treatment would cleave attB site and fusion tags from C terminal of the protein. The fusion tags could be easily removed from AcTEV treated protein solution by Ni chromatography, since the AcTEV treated protein would be in the flow through and the fusion tags would remain attached to the Ni via the His-tag. The AcTEV protease also had His-tag attached to it for easy removal along with the fusion tags.

Table 1 Lists of genes identified using computational analysis and primer design. Highlighted in blue are start codons and highlighted in orange is AcTEV site.

ID	Gene Number	Primer	Primer sequence
1	0284	F	ggggacaagtttgtacaaaaaagcaggctyygaggtgataagatgaataaaatttcaag
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggatttaacttcgcctgttttaggatc
2	0397	F	gggacaagtttgtacaaaaaagcaggctyygagggagtatatcatgagagaaaatttta
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggatttttctaattttggtgccttcg
3	0398	F	ggggacaagtttgtacaaaaaagcaggctyygaggagtgatagatatgaaatttaaaaagg
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggattttacgtgttcatgttttgaaa catc
4	0630	F	ggggacaagtttgtacaaaaaagcaggctyyaaggaggagtcccttgaaaaa
		R	ggggaccactttgtacaagaaagctgggygaaaacctgtattttcagggaatggatgaatgcatagctagaaact
5	0661	F	ggggacaagtttgtacaaaaaagcaggctyygtaaggaaatgtgggtaagggg
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggatttcgttattatgccttgattta
6	0720	F	ggggacaagtttgtacaaaaaagcaggctyygacagcaagatgaaaaaaactcttac
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggagtggatgtaattatattttcctggattt a
7	0754	F	ggggacaagtttgtacaaaaaagcaggctyyaggagaaataacagatgaaaaagaaattat
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggatactcgtggtgctggtaagcta
8	0755	F	ggggacaagtttgtacaaaaaagcaggctyyaagggagcataaacaaatgaaaag
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggagcgtttttcagttttctctatcttattatt
9	0964	F	ggggacaagtttgtacaaaaaagcaggctyyaagaggtgtttttatagaagaataaaatttttg
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggaataaatgttataatacgctctattgac cc
10	1055	F	ggggacaagtttgtacaaaaaagcaggctyyaggaggaattttctatgaaaaaagtaaac
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggatttattcgggattgtttgtttga
11	2368	F	ggggacaagtttgtacaaaaaagcaggctyyaggagctggtgatagatgtcg
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggacttttttatttagttcttttttttatatcatct gg
12	2373	F	ggggacaagtttgtacaaaaaagcaggctyyaggtggagaaatgatgattcattc
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggaccaggatttatttgatattccgttt
13	2374	F	ggggacaagtttgtacaaaaaagcaggctyyaggtggagaaatgatgatgca
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggagtttgactcataacttgtat cctcgc
14	2442	F	ggggacaagtttgtacaaaaaagcaggctyyaaggagtatgatagcgatgagaatattaa
		R	ggggaccactttgtacaagaaagctgggtygaaaacctgtattttcagggaaaagaaatcagatgggttaaattctt

Genomic DNA Isolation

A modified method from Current Protocol in Molecular Biology was used for preparation of genomic DNA from bacteria. Overnight culture (1.5 ml) of MSSA476 was centrifuged at 10,000 x g for 5 min. The pellet was resuspended in 567 µl TE buffer [10 mM Tris (hydroxymethyl) aminomethane HCl, pH 7.5 (Tris-HCl) with 1 mM ethylenediaminetetraacetic acid, pH 8.0 (EDTA)] and 30 mg/ml lysozyme was added and incubated for 20 min at room temperature. To the cell suspension, 30 µl of 10% sodium dodecyl sulfate (SDS) and 3 µl of 20 mg/ml proteinase K were added, mixed well, and incubated at 37˚C for 1 hr. Following incubation, 100 µl of 5 M NaCl was added and mixed thoroughly. To this, 80 µl Cetyl trimethylammonium bromide (CTAB) /NaCl solution (4.1 g NaCl, 10 g CTAB and make up volume 100 ml with water) was added, mixed vigorously and incubated at 65˚ for 10 min. The genomic DNA extraction was performed by adding 0.7 ml isoamyl alcohol:chloroform (24:1). The aqueous phase was separated and extraction was repeated once. Further extraction of genomic DNA was performed with 0.7 ml phenol:isoamyl alcohol:chloroform (25:24:1). Genomic DNA was precipitated with 420 µl isopropanol and washed with 75% ethanol. The pellet was air dried and resuspended in 50 µl TE buffer and stored at -20˚C.

Gateway Cloning and Protein Expression

The Gateway cloning kit with Clonase II was purchased from Invitrogen. The kit comes with all reagents, materials, protocols and cells required to perform two-step cloning via BP reaction and LR reaction to obtain Expression clones.

attB PCR Product

attB PCR product was synthesized for all 14 genes utilizing the following method. The isolated genomic DNA from MSSA476 was used as the template for all genes identified using computational analysis. Polymerase chain reaction (PCR) was set up using 2 µl genomic DNA and 1 µl each of the forward and reverse primers and was mixed with 8.5 µl nuclease free water. To this, 12.5 µl 2x GoTaq PCR master mix was added and vortex for 2 sec (Table 2). The PCR program was set up for denaturing, annealing and extension of target gene as listed in Table 3.

Table 2 PCR master mix composition and reaction setup details.

PCR master mix	**per reaction**
$MgCl_2$	4 µl
Go Taq green flexi color buffer	5 µl
PCR nucleotide mix	1 µl
GoTaq Polymerase	0.125 µl
Nuclease free water	2.4 µl
Total	12.5 µl

PCR reaction	
DNA template	2 µl
Forward Primer	1 µl
Reverse Primer	1 µl
Water	8.5 µl
PCR mix	12.5 µl
Total volume	25 µl

The PCR product for each gene was loaded onto a 1% agarose. The gel was prepared by adding 0.5 g agarose to 5 ml 1 x TAE buffer (40 mM Tris-acetate/1 mM EDTA, pH 8.5) and boiled in a microwave for 2 min or until completely dissolved. The agarose was allowed to rest on the bench top until the temperature was about 55˚C and then 2 µl of ethidium bromide was added, mixed by swirling and poured into gel cast. The gel cast had rubber edges that sealed them and had slots for placing the comb before pouring agarose gel into the cast. The gel was

allowed to solidify for 30 min. Once solidified, the comb was carefully removed, and the gel was submerged in 1 x TAE buffer. Agarose wells were loaded with 3 µl of the *att*B PCR product and 5 µl of 1 Kb DNA marker was also loaded in a well for size approximation. The gel tank was connected to a power supply and subjected to electrophoresis for 45 min at 100 V, constant voltage, followed by placing the gel on UV dock illuminated to detect DNA bands.

Table 3 Program set up to run PCR for all 14 genes.

PCR Program		
Step		
1	95˚C	2 min
2	94˚C	50 sec
3	55˚C	40 sec
4	68˚C	1 min
5	5 times to step 2	
6	95˚C	40 sec
7	65˚C	50 sec
8	68˚C	1 min
9	25 times to step 6	
10	68˚C	5 min
11	4˚C	0 min

Purifying attB Product

The attB product was purified to remove any contaminating primer dimers that may interfere with cloning. attB PCR product purification was performed using the following protocol from Invitrogen. In a microcentrifuge tube, 75 µl of TE buffer was added and 20 µl PCR reaction mix was added. To this, 50 µl of PEG (polyethylene glycol 8000 MW/30 mM MgCl2) was added, mixed vigorously and centrifuged for 15 min at 10,000 x g at room temperature. Very carefully, the supernatant was removed and the pellet was resuspended in 25 µl TE buffer. After purification, the *att*B product was checked for purity on 1% agarose gel.

BP Reaction

The *att*B PCR product was used to carry out the BP cloning reaction from Invitrogen Gateway cloning to obtain entry clones. The *att*B PCR product concentration was calculated using the formula provided by Invitrogen (formula not shown). About 50 femtomoles (fmol) of each purified *att*B PCR product of the genes of interest was used to perform BP recombination reactions in individual vials. In a microcentrifuge tube, 5 µl of *att*B PCR product was added plus 1 µl of pDONR vector (150 ng/µl), provided with the kit, and the reaction volume was increased to 8 µl with TE buffer. To the reaction mixture, 2 µl of BP clonase II enzyme mix was added, mixed by gently tapping the vial, and incubated in a 25˚C water bath for 18 hrs. Following incubation 1 µl proteinase K solution was added to each reaction and placed in a 37°C water bath for 10 min.

Transforming BP Reaction into Competent E. coli Cells

Following the BP reaction, DH5α competent *E. coli* cells were transformed using the protocol provided by Invitrogen. One vial of competent cells per gene was thawed on ice and 1 µl of the BP reaction was added and mixed by gently tapping the vial. The vials were incubated on ice for 30 min, the cells were subjected to heat shock at 42˚C for 30 sec, and returned to ice for another 2 min. To each vial, 250 µl of SOC media (Super Optimal broth with Catabolite repression) was added and incubated at 37˚C for 1 hr with horizontally shaking at 225 rpm. Luria-Bertani (LB) agar plates were prepared with 50 µg/ml kanamycin. Transformed cells were diluted 1:20 times with LB broth. Diluted cells (100 µl and 20 µl) are plated using the spread plate technique onto 2 LB agar plates containing 50 µg/ml kanamycin and incubated at 37˚C for overnight.

BP Clone Selection

Clones were selected based on the acquisition of kanamycin resistance and confirmed by plating the clones on LB agar plates with kanamycin and LB agar plate with chloramphenicol. Clones that were inhibited by chloramphenicol and were resistant to kanamycin were selected and colony PCR was performed. For colony PCR, the reaction setup and PCR program were the same as used when obtaining the *att*B PCR product (Table 2) but with modification to DNA template that was replaced with cells picked from a single kanamycin positive colony and mixed with the PCR reaction mixture. Following PCR, the reactions were loaded onto a 1% agarose gel as described earlier. Colonies with bands corresponding to cloned gene sizes were selected to perform the LR reaction to obtain expression clones.

Plasmid Isolation

Selected clones for each gene were grown overnight in separate 5 ml LB broth tubes containing 50 µg/ml kanamycin and subjected to plasmid isolation for each gene. The overnight cultures were used to isolate the plasmids that were later used to perform LR recombination reaction. Zippy plasmid isolation kit from Zymo research was used for plasmid isolation that provided all necessary buffers and supplies. To isolate plasmid, 600 µl of overnight culture was placed in a vial and 100 µl of 7 x lysis buffer was added and mixed by inverting. To this, 350 µl of neutralization buffer was added, mixed by inverting and centrifuged at 12,000 x g for 4 min. The supernatant was transferred into the spin column provided with the kit and placed into a collection tube and centrifuged at 10,000 x g for 15 sec. The flowthrough was discarded and 200 µl of eno-wash buffer was added to the spin column and centrifuged at 10,000 x g for 30 sec. To the spin column, 400 µl of Zippy wash buffer was added and centrifuged at 10,000 x g for 1 min. The column was transferred to a microcentrifuge tube, 30 µl of Zippy elution buffer was added

to the top of the column and allowed to sit for 5 min before centrifuging at 10,000 x g for 30 sec. Plasmid isolation was confirmed by loading samples onto a 1% agarose gel and detection of DNA bands corresponding to the plasmid size.

LR Reaction

The isolated BP plasmid was used to carry out the LR cloning reaction to obtain expression clones. Genes of interest cloned into the BP plasmids were used to perform the LR recombination reactions. The LR reaction was setup according to Invitrogen by placing 3-5 µl of BP plasmid (100 ng) in a vial and adding 1 µl of pET DEST42 vector (150 ng/µl) (destination vector) provided with the kit and adding 8 µl TE buffer. To the reaction, 2 µl of LR clonase II enzyme mixture was added to start the reaction, mixed by gently tapping the vial, and incubated in a 25˚C water bath for 18 hrs. Following incubation, 1 µl proteinase K solution was added to each reaction and placed in a 37˚C water bath for 10 min.

Transforming LR Reaction into Competent E. coli Cells

Following the LR reaction, DH5α competent *E. coli* cells were transformed using the same protocol as used in the "Transforming BP reaction into competent *E. coli* cells" (as described earlier). One vial of competent cells per gene was thawed on ice, 1 µl of the LR reaction added, and mixed by gently tapping the vial. The vials were incubated on ice for 30 min. The cells were subjected to heat shock at 42˚C for 30 sec and returned to the ice for another 2 min. To each vial, 250 µl of SOC media was added and incubated at 37˚C for 1 hr with horizontally shaking at 225 rpm. LB agar plates were prepared with 100 µg/ml ampicillin. Transformed cells were diluted 1:20 with LB broth. Diluted cells (100 µl and 20 µl) were plated

using the spread plate technique onto LB agar plates containing 100 µg/ml ampicillin and incubated at 37˚C overnight.

LR Clone Selection

Clones are selected based on the acquisition of ampicillin resistance, which was confirmed by plating the clones on LB agar plates with kanamycin and LB agar plates with ampicillin. Clones inhibited by kanamycin and resistant to ampicillin were selected and colony PCR was performed. For colony PCR, the reaction setup and PCR program were same as used when obtaining the *att*B PCR product (Table 2 with the exception that the DNA template was replaced with cells picked from a single ampicillin positive colony and added to the PCR reaction mix. Following PCR, the reactions were loaded onto a 1% agarose gel, as described earlier. Colonies with band sizes corresponding to cloned gene sizes were selected as described earlier and transformed into BL-21 *E. coli* competent cells for expression of proteins of interest.

BL-21 Transformation

Plasmids isolated from LR clones were transformed into BL-21 *E. coli* competent cells. BL-21 cells were purchased from Invitrogen and the transformation protocol was followed according to Invitrogen. BL-21 *E. coli* competent cells were thawed on ice and 2 µl of the LR plasmid added and mixed by gently tapping. The vials were incubated on ice for 30 min. The cells were subjected to heat shock at 42˚C for 30 sec and returned to the ice for another 2 min. To each vial, 250 µl of SOC media was added and incubated at 37˚C for 1 hr with horizontally shaking at 225 rpm. LB agar plates were prepared with 100 µg/ml ampicillin. Transformed cells are diluted 1:20 times with LB broth. 150 µl and 20 µl of the diluted cells are plated on 2 LB

agar plates (containing 100 µg/ ml ampicillin) using the spread plate technique and incubated at 37˚C for overnight.

Protein Expression

BL-21 *E. coli* clones carrying the gene of interest were selected based on the expression of proteins. BL-21 clones of the 14 genes were induced using the following protocol. BL-21 clones were inoculated into 5 ml LB broth containing 100 µg/ ml ampicillin and incubated overnight with shaking at 200 rpm. Following incubation, the cultures were diluted 1:20 in 10 ml LB broth with ampicillin and returned to incubation for another 2 hrs. When 0.4 optical density (OD) was achieved, the culture was subjected to induction by adding 0.2% arabinose and 1 mM isopropyl-β-D-thiogalactoside (IPTG) and incubated at 37˚C for overnight with shaking at 200 rpm. Following incubation, 1 ml of induced culture from all 14 genes was centrifuged and the pellets were resuspended in 100 µl SDS loading buffer and stored at -20˚C. These samples were thawed and heated at 90˚C for 10min. Samples were loaded onto an SDS gel along with protein markers for size reference and uninduced sample as a control.

MSSA476 Hypothetical Exported Proteins SAS2373 and SAS2374

Computational Analysis

Both SAS2373 and SAS2374 are hypothetical exported proteins. The number of nucleotides, amino acids and primer design for both genes are listed in Table 1 along with the data for the 12 other genes identified as exported proteins. In the MSSA476 genome, both genes are located adjacent to each other and the presence of hypothetical lipoproteins upstream of both genes, therefore, referred to as tandem lipoprotein on gene DB website.

Bacterial lipoproteins have characteristic, positively charged aa in the initial residues, a LIPOBOX about 40 aa upstream of start codon [64], and 7-22 aa residues between the positively charged residue and the LIPOBOX. This LIPOBOX consists of 4 amino acids with consensus [LVI] [ASTVI] [ASG] [C] combinations.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Genes SAS2373 and SAS2374 were expressed using IPTG and 0.2% arabinose for 4 hrs. At 1, 2, 3 and 4 hr time points, induced and uninduced samples (1 ml each) were removed and fractions were loaded onto an SDS gel to detect protein expression with induced samples and uninduced samples as control.

An SDS gel was prepared by putting together a gel caster consisting of 2 glass plates of different heights, spacers placed between glass plates, and sealing ends of both glass plates. A running gel was prepared as listed in Table 2 and poured carefully into the space between the glass plates, about 3/4 of the total gel area consisted of running gel. The running gel was overlaid with butanol in water solution; this step was to level the gel surface preventing a meniscus and help protein to separate in a straight line. After the running gel polymerized and the butanol/water was poured out, the gel was rinsed with distilled H_2O. The stacking gel was prepared as listed in Table 4. Place appropriate comb size between glass plates just enough to make well to hold 50 µl sample and pour stacking gel filling the remaining space between the glass plates and was allowed to polymerize for 30 min. comb was removed and the gel plates were attached to gel cassettee and placed in tank with running buffer (prepared as listed in Table 4).

Induced and uninduced sample pellets were dissolved in 100 µl SDS loading buffer (from Invitrogen) and heated to 90˚C for 10 min or until the pellet dissolved (reduced time of heating

or heating at lower temperature would result in thick gelatinous mass that was difficult to load or pipette). After heating pellets in SDS loading buffer, 20 µl of each sample was loaded into separate wells and protein marker with a range of 14 Kd-66 Kd was also loaded, as shown in Fig 2 (from Sigma Aldrich). The gel cassette was electrophoresed at 100 V, constant voltage and proteins were allowed to separate for 1 hr and 45 min and followed by Coomassie brilliant blue staining.

Table 4 Composition of running gel, stacking gel and buffers used to prepare the SDS gel.

12.5% Running gel	
30% acrylamide/ 0.8% bis	4.2 ml
H_2O	3.3 ml
1.5 M Tris, pH 8.8. 0.4% SDS	2.5 ml
TEMED	5.0 µl
10% Ammonium persulfate	50.0 µl
4% stacking gel	
30% acrylamide/0.8 % bis	1.3 ml
H_2O	6.1 ml
0.5 M Tris pH 6.8	2.5 ml
10% SDS	100.0 µl
TEMED	10.0 µl
10% Ammonium persulfate	50.0 µl
Running buffer	
Glycine	14.40 g
Tris	3.03 g
SDS	1.00 g
Staining solution	
Coomassie brilliant blue	0.25 g
Methanol	400.0 ml
Acetic acid	70.0 ml
H_2O	up to 1 L
Destaining solution	
Methanol	400.0 ml
Acetic acid	70.0 ml
H_2O	up to 1 L

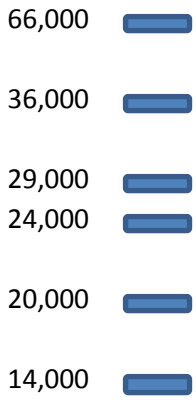


Figure 2 Protein marker purchased from Sigma-Aldrich M.W. 14,000 – 66,000.

Protein staining was achieved by using Coomassie brilliant blue stain (prepared as listed in Table 2) for 40 min and then washed with water and then destained for overnight in destaining solution (prepared as listed in Table 4). Followed by observing the gel under light to detect visible protein bands.

Western Blot

Overnight induced and uninduced fractions of both SAS2373 and SAS2374 were dissolved in SDS loading buffer, heated at 90˚C for 10 min, and loaded onto an SDS gel that was separated by electrophoresis for 1 hr and 45 min. Proteins were transferred from the unstained SDS gel onto a nitrocellulose membrane overnight in transfer buffer (25 mM Tris/ 192 mM Glycine/ 20% methanol/ up to 1L, pH 8.3) at 100 milliamps power and then increased the next day to 200 milliamps for 1 hr. The nitrocellulose paper was subjected to a series of wash steps, followed by detection with anti-His antibodies. The required buffers and solutions for washing were prepared as listed in Table 5 and washing was performed in following order. The nitrocellulose membrane was washed with washing solution I for 15 min twice, followed by a 1 hr incubation in blocking solution, then washed twice in washing solution II for 10 min each and

in 15 ml washing solution I for 10 min. The membrane was then soaked in 10 ml His-Tag monoclonal antibody for 1 hr and washed twice with washing solution II for 10 min. It was then washed with 15 ml washing solution I for 10 min, followed by soaking the membrane in 8ml goat anti-mouse antibody for 1 hr and then washing 5 times with washing solution II for 10 min each. The developing solution was prepared in 15 ml 1 x AP buffer (Alkaline phosphatase buffer) and 60 µl NBT (83 mg nitroblue tetrazolium/ ml in 70% dimethylformamide)/ 60 µl BCIP (42 mg 5-bromo-4-chloro-3-indolyl phosphate /ml in 100% dimethylformamide) solutions were added. The membrane was placed in clean tray and exposed to the developing solution. Visible signal was observed for 2-10 min. Any signal after 10 min is considered a false positive. The reaction was stopped by washing the membrane with deionized water.

Protein Purification

Protein purification was achieved by inducing the gene of interest to produce high amounts followed by Ni column chromatography. BL-21 *E. coli* expression clones for SAS2373 and SAS2374 were induced and processed for purification of protein. For each protein, the BL-21 *E. coli* clone was inoculated into a 250 ml conical flask containing 50 ml LB broth with 100 µg/ml ampicillin. The culture was grown overnight at 37˚C on a shaker at 200 RPM. Following incubation, the culture was diluted 1:20 with LB broth/ampicillin (50 ml overnight culture and 950 ml LB amp broth) and returned to shaker for another 2 hrs till 0.4 OD is achieved. The cells were induced by adding 1 mM IPTG/ml and 0.2% L-arabinose and returned to the shaker for overnight induction of protein.

After the protein was induced, the broth culture was subjected to series of purification steps to obtain a cell extract. A His-bind purification kit from Novagen was used to purify the

proteins from inclusion bodies. The buffers required for protein purification were prepared as listed in Table 6.

Table 5 Composition of buffers and solutions used to perform Western blot.

Blocking solution	
BSA	1.00 g
20 x TBS	1.65 ml
H_2O	33.30 ml
Washing solution I	
20 x TBS	3 ml
H_2O	60 ml
Washing solution II	
10 x TSBTT	18 ml
H_2O	180 ml
His Tag Antibody	
Monoclonal antibody	5 µl
Blocking solution	10 ml
Goat anti-mouse Antibody	
Antibody	1.6 µl
Blocking solution	8.0 ml

The overnight-induced culture was centrifuged at 10,000 x g for 15 min. The pellet was dissolved in 40 ml binding buffer/100 ml culture and subjected to sonication on ice at 30 amplitude pulse for 6 min, then cooled for 5 min, and repeated 3 times. The homogenized mixture was centrifugation at 5,000 x g for 15 min. The pellet was dissolved in 20 ml/100 ml culture, sonicated and centrifuged again. The pellet was dissolved in 5 ml/100 ml culture volume solubilizing buffer (1 x binding buffer/6 M Urea) and placed on ice for 1 hr. This was followed by centrifugation at 16,000 x g for 30 min, and the supernatant was collected and filtered through 0.45 µm syringe filter. This filtered cell extract consisted of the protein of interest along with many other BL-21 *E. coli* proteins as contaminants. The protein of interest was isolated by passing the cell lysate through a Ni chromatography column.

Table 6 Composition of buffers used for performing protein purification.

8x Binding buffer	
NaCl	4 M
Tris- HCl	160 mM
Imidazole	40 mM
pH	7.9
8x Wash buffer	
NaCl	4 M
Tris-HCl	160 mM
Imidazole	480 mM
pH	7.9
4x Elute buffer	
NaCl	2 M
Tris-HCl	80 mM
Imidazole	4 M
pH	7.9

The Ni column was prepared according to the Novagen His Bind protein purification manual. Pre-charged Ni- resin was placed in a column to a final bed volume of 500 µl. The resin surface was made even by placing a syringe frit. The resin was washed with 1.5 ml deionized water and then with 1.5 ml of 1 x binding buffer. Binding buffer was allowed to flow through until it reached the top of resin bed. While the mouth of the column was closed, very slowly the cell extract was poured onto the surface of the column and allowed to stand for 15 min before allowing the cell extract to flow through. The Ni-resin was washed with 5 ml 1 x binding buffer and then twice with 3 ml each of 1 x wash buffer (Modified from Novagen protocol). When the wash buffer drops to the top of resin bed the protein was eluted using 3 ml 1 x elution buffer. The eluted protein was collected in 1 ml fractions and stored at -20˚C for further use. The eluted fractions were loaded onto an SDS gel and checked for purity.

Dialysis and AcTEV Protease Treatment

Dialysis was performed to gradually remove imidazole from the purified protein solution over a period of 18 hrs. Purified protein was carefully placed in dialysis membrane tubing with cutoff molecular weight of 14,000 daltons, meaning it would retain proteins larger than 14,000 daltons. The tubing was clamped making sure not to pull the membrane or accidentally enlarge the pore size which may lead to protein leakage. The tubing was placed in three buffer solutions varying in concentration as listed in Table 5. Dialysis was performed with solution I and II (prepared as listed in Table 7) for 1.5 hrs each and solution III (prepared as listed in Table 7) for overnight at 4˚C. After dialysis the protein was carefully removed from the tubing and transferred to sterile vials after passing through 0.45 µm syringe filter to remove any contaminants.

AcTEV protease: Protease treatment removes the additional amino acids attached to the proteins at the C terminal end. The proteins identified by computational analysis are hypothetical exported proteins and the function is unknown. Very little is known about these proteins, and therefore removing the fusion tags or any additional amino acids from C terminal that were added as a result of the cloning process, will increase our chances of the protein being folded correctly and thereby making an active product.

Protease treatment was performed following AcTEV Protease protocol from Invitrogen. To 100 µl of purified protein was added 7.5 µl of 20 x TEV buffer (1 M Tris-HCl, pH 8 and 10 mM EDTA), 1.5 µl of 0.1 M DTT, and 1 µl of AcTEV protease (10 units) and the volume made up to 150 µl by adding 40 µl sterile deionized water. This reaction was incubated at 30˚C for 6 hrs. At 2 hrs intervals from time zero (2, 4 and 6 hrs) 30 µl of the reaction was removed into separate vials and mixed with 30 µl of SDS loading buffer and heated for 10 min at 90˚C. Each

of reaction samples were loaded onto an SDS gel along with protein markers and untreated purified protein as control. After running the gel for 1 hr and 45 min, it was stained with SDS staining solution (Table 2) for 40 min, then destained for overnight in destaining solution (Table 2).

Table 7 Composition of dialysis buffers solutions.

Solution I	
Tris	20 mM
NaCl	400 mM
Imidazole	1 M
pH	7.9
Solution II	
Tris	20 mM
NaCl	200 mM
Imidazole	250 mM
pH	7.9
Solution III	
Tris	20 mM
NaCl	50 mM
pH	7.9

Microbial Killing Assay

Proteins can be involved in competing against normal flora of skin to colonize or cause infection. To determine if the proteins of interest were involved in inhibiting growth or killing normal flora of skin, the normal flora were directly exposed to purified protein. The overnight culture of *Micrococcus luteus*, *Staphylococcus epidermidis*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* were grown in 5 ml LB broth medium. Broth culture (100 µl) was spread on an LB agar plate with a spreader to make even lawn of bacteria. Purified protein (25 µl and 100 µl) were placed directly at two different spots; a 3rd spot was loaded with sterile dialysis solution III as a control. All plates were incubated at 37˚ for overnight growth and were observed for the zone of inhibitions.

Caenorhabditis Elegans

Nematode growth medium (NGM) agar plates was prepared with 1 mM/ml IPTG and 0.2 % arabinose and a 2nd set of plates was prepared without IPTG and arabinose. All protocols for growth and maintenance of *C. elegans* were used from (CGC). *C. elegans* feeds on *E. coli* stain OP50 which was also purchased from CGC. Overnight cultures of *E. coli* OP50 were grown at 37˚C/200 rpm in LB broth and a lawn was made in the center of NGM agar petri plate and incubated at 37˚C for overnight growth. Individual *C. elegans* animals were picked from the source plate with the help of a hook shaped loop, placed in the center of OP50 lawn and incubated at room temperature for 3-4 days. On day 2, BL-21 *E. coli* SAS2373 and SAS2374 clones were inoculated into LB broth with ampicillin and incubated at 37˚C/ 200 rpm for overnight growth. On day 3, the overnight cultures were diluted 1:20 in LB broth with ampicillin and incubated for 2 hrs at 37˚C/200 rpm. After 2 hrs, when 0.4 OD was achieved the bacterial culture was placed on NGM agar plates with and without IPTG and arabinose, using a spreader. These plates were incubated overnight at 37˚C. On day 4, adult *C. elegans* were transferred onto the NGM agar plated with the BL-21 clones of SAS2373 and SAS2374 lawns. Taking care to use equal numbers, the animals were plated on each plate and incubated at room temperature for 3-4 days.

Co-Culture

Co-culturing Keratinocytes with MSSA476

Keratinocytes were purchased from Zenbio, which also provided necessary instruction for growth and maintenance. The cryopreserved keratinocyte vial was thawed in a 37˚C water bath for 2 min and the entire contents of the vial poured into 9 ml keratinocyte growth media

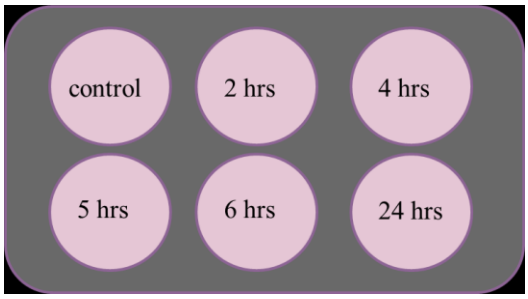
(KM). Cells were centrifuged at 400 x g for 10 min at 20˚C. The supernatant was removed and the cells were re-suspended in 3 ml KM media and counted using a haemocytometer. Cells (5000 cells/cm^2) were added to 5 ml KM media, placed into 25 cm^2 flasks and incubated at 37˚C with 5% CO_2 for 24 hrs. Media was replaced with fresh media after 24 hrs and then every 2 days.

When keratinocytes had reached 80% confluent, cells were split into 6 well plates. Before splitting cells all solution and media were warmed to 37˚C. KM media from the flasks was removed and the cells washed with phosphate buffered saline (PBS) three times to make sure all traces of media were removed from the flask. This was followed by the addition of 1 ml 0.25% trypsin to the flask and placed back in 37˚C incubator for 2 min. Cell detachment was checked before moving on. Once the cells were detached, 1ml of trypsin inhibitor was added to stop the reaction of the trypsin and the cells were collected into centrifuge tubes. The flask was washed with 2 ml KM media and added to the centrifuge tube to ensure all cells were harvested from the flask. The cells were centrifuged at 400 x g at 20˚C for 5 min. The supernatant was removed and the cells were re-suspended into 2 ml KM media. Cells (5000 cells/cm^2) were placed into 6 well plates and incubated at 37˚C with 5% CO_2. After 24 hrs incubation, the KM media was replaced and the plate returned for further incubation; subsequently, the media was replaced every 2 days until the cultures achieved 80% confluence.

Keratinocytes Co-Culture Set Up

On the day of the co-culture experiments, 1 ml of overnight MSSA476 culture was pelleted and washed with sterile deionized water and resuspended in 1 ml of LB broth which was diluted 10^{-6} times in sterile deionized water. The KM media from the 6-well keratinocyte media plate was replaced with antibiotic free KM media to ensure that the antibiotic does not kill the bacterial cells. After replacing the media, each well was inoculated with 25 µl of diluted

MSSA476 culture and incubated for 24 hrs (Fig 3a) and 25 µl was also plated on an LB agar plate and incubated at 37˚C. Colonies were counted the next day. Alongside the keratinocytes plate, another 6-well plate was prepared without keratinocytes (Fig 3b) as the control with antibiotic free media and MSSA476. Cells were harvested from the control plate and the keratinocytes 6- well plate at 2, 4, 5, 6 and 24 hr time points and subjected to RNA isolation as soon as the cells were harvested (samples labeled as listed in Table 8)

a. 

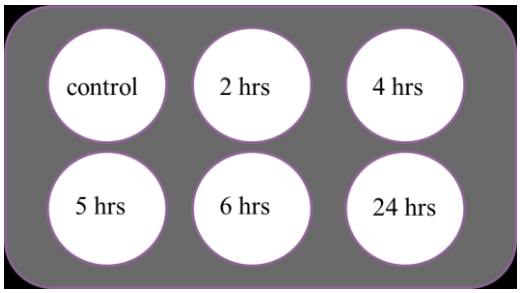
b. 

Figure 3 Co-culture and control plate set up.

Six well plates inoculated with MSSA 476 in antibiotic free KM media. One plate contains 80% confluent keratinocytes (Fig 3a) and the 2nd plate is the control without keratinocytes (Fig 3b). The control well in 10a is a negative control, without MSSA 476 and control well in 10b is without both keratinocytes and MSSA476, thus, a sterility control.

Co-Culturing HEK (human embryonic kidney cell line) with MSSA476

Cryopreserved HEK cells were thawed at room temperature and poured into 100 mm tissue culture petri plate and 9 ml of Dulbecco's Modified Eagle Medium (DMEM) was added. The media was swirled to distribute cells equally and incubated at 37˚C with 5% CO2. It takes 2-3 days for cells to grow to 100% confluence depending on the initial inoculum. After the cells are 90-100 % confluent, they were split by washing twice with PBS and then 2 ml of 0.05% trypsin was added and swirled. After 2 min, when the cells had detached, trypsin was collected in

centrifuge tubes and the plate was washed with 2 ml DMEM and collected in the same centrifuge tube. The tube was centrifuged at 2,000 x g for 5 min and the supernatant carefully removed. The pellet was suspended in 10 ml media and 10,000 cells/well/6-well plate were seeded in 2 ml DMEM/well and incubated at 37˚C with 5% CO_2 for 2-3 days until the cells were 90-100 % confluent.

Table 8 List of isolated RNA and synthesized cDNA sample ID's.

Each sample ID is listed next to setup of the experiment and sample hr.

ID	Sample
K-2	Keratinocytes and MSSA476 Co-cultured for 2 hours
K-4	Keratinocytes and MSSA476 Co-cultured for 4 hours
K-5	Keratinocytes and MSSA476 Co-cultured for 5 hours
K-6	Keratinocytes and MSSA476 Co-cultured for 6 hours
K-24	Keratinocytes and MSSA476 Co-cultured for 24 hours
Km-2	KM media and MSSA 476 incubated for 2 hours
Km-4	KM media and MSSA 476 incubated for 4 hours
Km-5	KM media and MSSA 476 incubated for 5 hours
Km-6	KM media and MSSA 476 incubated for 6 hours
Km-24	KM media and MSSA 476 incubated for 24 hours
H-2	HEK and MSSA476 Co-cultured for 2 hours
H-4	HEK and MSSA476 Co-cultured for 4 hours
H-5	HEK and MSSA476 Co-cultured for 5 hours
H-6	HEK and MSSA476 Co-cultured for 6 hours
H-24	HEK and MSSA476 Co-cultured for 24 hours
D-2	DMEM and MSSA 476 incubated for 2 hours
D-4	DMEM and MSSA 476 incubated for 4 hours
D-5	DMEM and MSSA 476 incubated for 5 hours
D-6	DMEM and MSSA 476 incubated for 6 hours
D-24	DMEM and MSSA 476 incubated for 24 hours

When the cells were confluent, the co-culture was set up exactly as described for keratinocytes co-culture, with MSSA 476 cells diluted 10^{-6} times and using antibiotic free DMEM media before inoculating wells with MSSA476.

RNA Isolation and DNase Treatment

RNA was isolated from harvested co-culture samples using Ribozol RNA extraction reagent from Amresco. The protocol was followed as described by Amresco with modification to harvested cells as following. After the cells were harvested by centrifugation at 10,000 x g for 2 min, the pellet was suspended in 800 µl LB broth and 200 µl phenol/ethanol mix (5% phenol/95% ethanol, pH < 5). This step is critical to maintain RNA stability [65]. Samples were then mixed by vortexing and centrifuged at 10,000 x g for 5min. The pellet was suspended in 100 µl lysozyme solution (3 mg/ml) and incubated at 37˚C for 20 min [66] followed by centrifugation at 10,000 x g for 2 min and removal of the supernatant. The following procedure is as described by Amresco Ribozol RNA Extraction protocol for isolation from bacterial cells. The pellet was dissolved in 1 ml ribozol by pipetting and incubated for 10 min at room temperature. To the ribozol suspension, 200 µl chloroform was added, mixed vigorously for 15 seconds, and incubated for another 3 min at room temperature. The cell suspension was centrifuged at 12,000 x g at 4˚C for 15 min and the colorless upper layer of the phase was carefully collected, placed in a separate vial, and 0.5 ml of isopropanol was added and incubated at room temperature for 10 min to precipitate RNA. The sample was centrifuged at 12,000 x g for 10 min at 4˚C, the supernatant removed, and the pellet washed with 1 ml of 75% ethanol by vortexing, and centrifuged at 7,500 x g for 5 min at 4˚C. The pellet was resuspended in 40 µl RNase free water, mixed by pipetting and incubated for 10 min at 60˚C. After incubation the isolated RNA was stored at -70˚C until further use.

DNase Treatment

DNase treatment was performed as described by Invitrogen. The DNase reaction buffer (200 mM Tris (pH 8.4), 20 mM MgCl2 and 500 mM KCl) and 25 mM EDTA solution were prepared using RNase free solutions. The isolated RNA samples were subjected to DNase treatment by adding 8 µl RNA sample to 1 µl DNase reaction buffer, 1 µl of DNase I and incubating at room temperature for 15 min. DNase I was inactivated by adding 1 µl of 25 mM EDTA solution and incubating for 10 min at 65˚C. This protocol was utilized for DNase treatment of all RNA samples isolated at different time points.

Reverse Transcription Polymerase Chain Reaction (RT PCR)

For RT PCR, cDNA was prepared using an iScript cDNA synthesis kit from Bio-Rad. The reaction was set up as described by Bio-Rad, as listed in Table 9, for all RNA samples the cDNA obtained was diluted 1:20 and stored at -20˚C until further use.

Table 9 RT reaction set up and RT PCR program.

Reaction	
5x iScript reaction mix	4 µl
iScript reverse transcriptase	1 µl
Nuclease-free water	9 µl
DNase treated, RNA template	6 µl
Total volume	20 µl

RT PCR program	
25˚C	5 min
42˚C	30 min
85˚C	5 min
4˚C	0 min

Polymerase Chain Reaction (PCR)

Each cDNA sample was used as the DNA template and PCR was performed for the twenty genes listed in Table 10. The PCR reaction was set up as listed in Table 3. After the reaction was complete the samples were loaded onto a 1% agarose gel prepared with 1 x TAE buffer and 2 µl ethidium bromide. After running samples for 45 min, the gel was placed on a UV dock and bands corresponding to the respective genes were identified.

Quantitative Polymerase Chain Reaction (qPCR)

Samples that were positive using PCR were subjected to qPCR. The GoTaq qPCR master mix was purchased from Promega. All reactions and setup were followed as provided by Promega. A 25 µl reaction was set up in the reaction plate in duplicates. To set up the reaction, 2 µl of the template were placed in a reaction plate on ice. In separate vials, 12.5 µl of 2 x qPCR master mix, 1 µl each of the forward and reverse primers and 9.5 µl nuclease free water was added per reaction and mixed gently. The reaction mixture was pipetted from the vials and transferred into reaction plate. When all transfers are made the reaction plate was centrifuged for 1 min at 1,000 x g.

The reaction plate was placed in the qPCR machine (MC3000p machine). The program was set for SYBR Green dye, using the 40 cycle program, as described in Fig 4. After the reaction was complete the results were analyzed using the MxPro qPCR Software.

To determine if the protein had elevated or decreased expression, 16S rRNA was used as the reference gene. Using the Livak methods [67], the fold change in protein expression was determined in comparison with reference gene. To perform this, the 16S rRNA primers were purchased from Invitrogen and PCR was performed against cDNA samples from keratinocyte

co-culture and HEK co-culture. The results of the reactions were loaded onto 1% agarose gels and DNA amplification bands were detected.

Table 10 List of genes used for PCR and qPCR using cDNA as template.

S.no	Gene Number
1	SAS0284
2	SAS0397
3	SAS0398
4	SAS0630
5	SAS0661
6	SAS0720
7	SAS0754
8	SAS0755
9	SAS0964
10	SAS1055
11	SAS2368
12	SAS2373
13	SAS2374
14	SAS2442
15	SAS2189
16	SAS1738
17	SAS0760
18	SAS1678
19	SAS0347
20	SAS1089

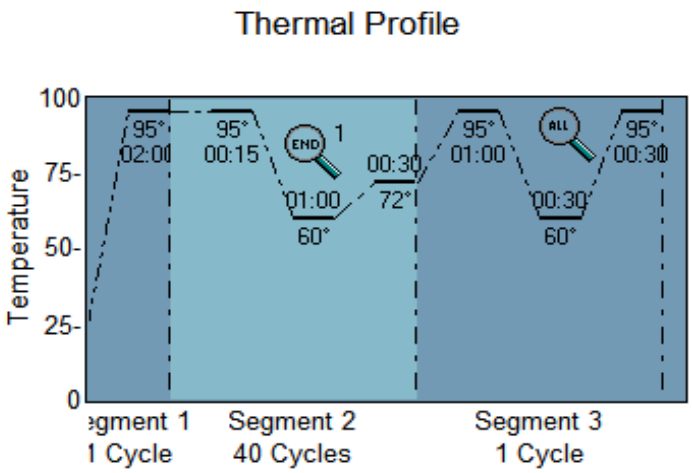


Figure 4 Thermal profile set up for running qPCR.

Cycles were modified from the original software program by increasing the annealing temperature to 60ºC, instead of 55 ºC in both 2 and 3 segments.

CHAPTER 3

RESULTS

Computational Analysis

Gene Database revealed 363 genes described as hypothetical proteins in the *S. aureus* MSSA476 genome. Using computational analysis, 47 genes were identified as hypothetical exported proteins. Each of the 47 genes had signal peptide, cleavage site and was non-cytoplasmic in nature, meaning they were secreted outside the cell wall and were negative for transmembrane domains. All 47 genes were further classified based on percent identity to HA-MRSA or CA-MRSA. Twenty four genes (listed in Table 11) were found to have 100% identity with CA-MRSA proteins (Table 12). Eight of these had recently been cloned in our laboratory. Remaining 16 genes primers were designed using Invitrogen primer design software to perform gateway cloning.

Primer Design

Primers were designed for 16 genes using Invitrogen primer design tool, but successful primer designs were only obtained for 14 genes. Primers could not be designed for the remaining two genes, SAS1055 and SAS2525, using the Invitrogen primer design program with varying parameters.

Table 11 List of genes identified as hypothetical exported proteins with 100 % identity towards CA-MRSA.

Table lists gene number, number of amino acids, whether the protein is secreted, cleavage prediction of the signal peptide and percent identity towards HA-MRSA and other strains of *S. aureus*.

Gene Number	Amino acids	Nucleotides	Secreted	Signal Peptide prediction	Cleavage prediction	site of cleavage	% identity
SAS0284	296	891	yes	1	0.992	31-32	99% HA MRSA
SAS0397	500	1503	yes	0.903	0.863	32-33	91% HA MRSA
SAS0398	102	309	yes	1	0.936	29-30	99% HA MRSA, Newman
SAS0630	265	798	yes	0.999	0.997	25-26	99% HA MRSA
SAS0661	154	465	yes	0.998	0.635	37-38	99% HA MRSA
SAS0720	279	840	yes	0.999	0.996	24-25	99% HA MRSA
SAS0754	340	1023	yes	0.971	0.971	26-27	99% HA MRSA, Lab
SAS0755	173	522	yes	0.522	0.481	26-27	98% HA MRSA
SAS0964	93	282	yes	0.998	0.524	23-24	99% HA MRSA
SAS1055	308	927	yes	0.734	0.734	32-33	99% HA MRSA, Newman
SAS1088	105	318	yes	1	0.997	29-30	99% HA MRSA
SAS2368	136	411	yes	0.85	0.784	21-22	97% VISA
SAS2373	271	816	yes	0.949	0.511	30-31	88% HA MRSA
SAS2374	272	819	yes	0.987	0.633	30-31	74% HA MRSA
SAS2442	66	201	yes	0.544	0.504	23-24	99% HA MRSA
SAS2525	157	474	yes	0.976	0.904	28-29	96% Col, Lab
SAS2189	269	810	yes	1	0.999	27-28	99% HA MRSA

Table 12 List of *S. aureus* strains with whole genome sequence available for computational analysis.

This list of strains was used to identify percent identity of MSSA476 hypothetical genes to CA MRSA and HA MRSA strains and then categorized as hypothetical exported proteins with 100% identity towards CA MRSA.

CA-MRSA
MW2
MSSA476
USA300/TCH_1516
USA300/TCH_FRP3757

HA-MRSA
N315 (MRSA/VSSA)
Mu50 (MRSA/VISA)
Mu3 (MRSA/VISA)
JH1- JH9 (MRSA/VRSA)
MRSA252 (MRSA)
COL (MRSA)

Other Strains
NCTC 8325
Newman
ED98
RF122

Gateway Cloning

PCR amplification of 14 genes was performed simultaneously using MSSA 476 genomic DNA (Fig 5) as a template and the amplified products were detected on 1% agarose gel (Fig 6). One gene, SAS0397, was not amplified due to its size (1503 base pairs) but was successfully amplified later by increasing the amount of nucleotide master mix in the PCR reaction set up.

BP Reaction

The BP reaction was successfully performed with 14 genes and transformed into competent cells. The entry clones for each gene were capable of growing on 50 µg/ml kanamycin and were inhibited by 30 µg/ml chloramphenicol. Entry clones capable of resisting kanamycin and losing chloramphenicol resistance indicated successful transformation. This was also confirmed by colony PCR reaction showing DNA fragment size corresponding to the respective genes on a 1% agarose gel (Fig 7).

LR Reaction

The LR reaction was performed following BP plasmid isolation, transformed into competent cells and plated onto LB agar plates containing 100 µg/ ml ampicillin. With the LR recombination reaction, clones with successful transformation will acquire ampicillin resistance and lose resistance to kanamycin. Therefore, expression clones capable of growing on ampicillin media plates and inhibited on kanamycin media plates were selected and subjected to colony PCR. Colony PCR reactions showing fragment size corresponding to respective gene size were selected for BL-21 transformation.

All but one gene were successfully cloned. Gene SAS2368 could not be cloned into the expression vector. At least five trials were made to clone this gene but all failed. Modification to methods by varying the amount of plasmid DNA, increasing the amount of enzyme, increasing the amount of LR reaction mixture, and the one-tube gateway cloning method were employed, but no modification aided in achieving LR clones.

Typical results following the LR reaction showed SAS2368 clones were transformed and plated. Very few colonies grew and all the colonies grew on both kanamycin and ampicillin

containing plates. The colonies were isolated and grown on fresh plate but all colonies came out negative for PCR amplification of the gene.

BL-21 Transformation and Protein Expression

Plasmid isolated from the isolated LR expression clones was transformed into BL-21 *E. coli* competent cells. The Bl-21 clones with successful transformation were identified by colony PCR.

All 13 genes induced overnight were loaded onto an SDS gel and stained but the bands for induced protein were difficult to identify for all genes. The uninduced and induced protein samples consisted of proteins associated with BL-21 *E. coli* cells as well making it difficult to detect the induced proteins on SDS gel.

SAS2373 and SAS2374 Computational Analysis

Both genes SAS2373 and SAS2374 were negative for the presence of the predicted LIPOBOX. The consensus sequence for LIPOBOX was not detected at 40 aa from start codon. To be sure, the LIPOBOX consensus sequence was checked by sequence matching from the start codon to 60aa. The LIPOBOX was not identified in either gene.

From protein BLAST search of SAS2373, we observed 100% identity of this protein only to CA-MRSA and 88% identity towards HA-MRSA. A BLAST of SAS2374 showed 100% identity towards MSSA476, 94% identity towards CA-MRSA and 74% identity towards HA-MRSA.

SDS and Western Blot

Genes *sas*2373 and *sas*2374 were induced for different time points, the samples were collected and loaded onto SDS gel and compared to uninduced samples (Fig 8). The native size

of each protein is about 31.5 Kdal. The protein fragments identified with induced samples were approximately 36 Kdal in size. The difference in size is due to the C terminal addition of the attB2 site (8 aa), AcTEV protease site (7 aa), V5 epitope (12 aa) and His tag (6 aa); all together adding to about 4 Kdal to the protein size.

Western blot was performed with the SAS2373 and SAS2374 protein samples. With the addition of substrate, dark purple color was observed within 2 min for both the proteins (Fig 9). SAS2374 revealed 2 bands of similar size; the top band corresponds to the size band that was expected while the lower band corresponds to the protein that is slightly smaller, possibly having cleaved the 5 aa N terminal cytoplasmic region. If the signal peptide was cleaved off which is 24 aa, the size difference would have been at least 3.5 Kdal and if the fusion tags (26 aa) were cleaved off the size difference would be greater and would not be detected on Western blot, as western blot antibodies targeted the His-tag.

With SAS2373 Western blot, 2 bands were identified in single lane, but with a greater size difference than was seen in SAS2374. The signal peptide is about 24 aa which corresponds to the size difference of about 4 Kd. Indicating the signal peptide was cleaved off in case of SAS2373.

Protein Purification

SDS gels revealed that we could isolate large amounts of protein from overnight induction. Therefore, overnight induced culture was used to purify both proteins (SAS2373 and SAS2374). The cell lysate was passed through the Ni column and eluted using imidazole. The eluted protein was filtered and loaded onto SDS gel as described in methods section. Purified protein showed a band size corresponding to the size of the protein with additional 4 Kdal fusion tags (Fig 8). In addition, some protein contaminants were also identified in low amounts. At this

point protein contaminants were not an issue because, after dialysis, AcTEV protease treatment required passing the protein through the Ni column again, which would help in removing protein contaminants.

Dialysis and AcTEV Proteases

Following protein purification, dialysis was performed with three buffer changes for both proteins. Dialyzed proteins were subjected to AcTEV proteases and the 2, 4 and 6 hr treated samples were loaded onto SDS gel along with marker and untreated samples, electrophoresed and stained.

When comparing the treated samples with untreated samples, it was observed that not all of the protein was cleaved by AcTEV protease. Only a small amount of protein was cleaved, identified as an additional band corresponding to cleaved protein size in the treated protein lanes, and absent in untreated sample lane (Fig 13).

Microbial Killing Assay

Dialyzed protein was directly placed on top of bacterial lawn. Normal flora of skin exposed to protein, SAS2373 or SAS2374 plus dialysis buffer, showed no inhibition of growth or zone of inhibition around the protein spots placed on the bacterial lawn.

Caenorhabditis Elegans

Caenorhabditis elegans nematodes were exposed to proteins (SAS2373 and SAS2374) on NGM plates with and without IPTG plus arabinose. The nematodes did not exhibit any bordering or reduction in number (Fig 10 & 11). However, an unexpected observation was made; on NGM plates with IPTG plus arabinose, the nematode showed enhanced growth with a lawn of Bl-21 or OP50 *E. coli* culture (Fig 12).

Co-Culture

Co-culture of MSSA476 with keratinocytes for 24 hrs did not show adverse effect on keratinocytes cells. It was observed that keratinocytes were not rounded or detached from the plates. Large colonies of keratinocytes remained looking healthy after 24 hrs whereas a few cell colonies had rounded and looked affected by MSSA476.

In keratinocytes growth media, the bacterial cell growth was checked after 24 hrs incubation with keratinocytes by plating on mannitol salt agar plate (MSA) and counting the colonies. The cell count was determined to be 10^{-7} cells/ ml equal to one log less than MSSA476 grown in LB broth medium for the same period of time.

RNA Isolation and DNase Treatment

RNA was isolated and PCR performed to make sure RNA samples were not contaminated with DNA because it will give false positive results. RNA samples were subjected to PCR with SAS2373 gene primers and loaded onto a 1% agarose gel. Bands corresponding to the correct gene size were identified confirming that the RNA samples were contaminated with DNA. Following DNase treatment, PCR was repeated with the same samples to confirm that the RNA was free of DNA (Fig 14)

RT PCR

RT-PCR performed with cDNA against twenty genes (genes listed in Table 10) showed positive expression of mRNA (Table 13) in presence of keratinocytes on a 1% agarose gel (Fig 15-17). PCR performed with HEK samples and DMEM media control samples (Table 8) were all negative for amplification of all twenty genes; no amplification bands were detected on the agarose gel.

Table 13 List of positively amplified genes using RT PCR.

The table shows co-culture samples and gene IDs (as listed in Table 10) which were positively amplified by that sample's cDNA as template. All sample descriptions are listed in Table 8 and gene IDs are listed in Table 1.

Samples	PCR with cDNA as template
K-5	4,5,12,13,16, & 17
KM-5	5,6,9,11,12, & 13
K- 6	12, 13 & 17
KM - 6	Negative
K-24	4, 11, & 12
Km-24	11 & 12

Quantitative PCR

Quantitative PCR was performed with cDNA samples that showed positive amplification with PCR (listed in Table 13), using 16S rRNA as the control. The C(T) values for the genes are listed in Table 14. The C(T) value obtained should be less than 35 with a 40 cycle reaction. Therefore, C(T) values more than 35 are highlight in purple (Table 14) and were not included in analysis. The only C(T) values considered were less than 35 and are highlighted in blue (Table 14). The dissociation curve and amplification plots for gene SAS2373 and SAS2374 showed positive amplification (Fig 18 & Fig 19).

Gene ID 12 and 13 showed expression with keratinocytes and with KM media only at 5 and 6 hrs. At 24 hrs, gene 13 showed expression in keratinocytes and in KM media only samples (Fig 21), whereas gene 12 showed expression in only in the presence of keratinocytes at 24 hr and was negative for expression in KM media alone (Fig 20).

Table 14 Shows C(T) values obtained by qPCR of co-cultured samples and control sample cDNA.

Genes that showed positive amplification with PCR (Table 13) were subjected to qPCR. Table 14 shows C(T) values for the qPCR. Sample K were co-cultured keratinocytes with MSSA476 and Km samples were keratinocyte growth media only inoculated with MSSA476 (control). Any reading above 35 C(T) value (highlighted in purple) was considered a false positive and not included in analysis. C(T) values below 35 were ruled as positive amplification and used for analysis using the Livak method [67] (Table 15).

Samples/ gene #	**4**	**5**	**9**	**11**	**12**	**13**	**16**	**17**	**16S**
5K	35.46	No Ct	38.4	No C	No Ct	No CT	37.95	38.92	26.92
5Km	38.58	No Ct	38.35	No Ct	No Ct	No CT	No Ct	No Ct	27.26
6K	No Ct	No Ct	No Ct	No Ct	30.25	27.7	No Ct	37.07	27.31
6Km	No Ct	No Ct	No Ct	No Ct	36.85	29.4	No Ct	No Ct	27.12
24K	36.57	38.56	No Ct	36.07	26.93	25.88	No Ct	37.55	22.66
24Km	37.6	37.25	No Ct	39.65	No Ct	31.43	No Ct	No Ct	21.7

The Livak method was used to determine fold increase or decrease in expression of protein with C(T) values less than 35 (Fig 22). The results are listed in Table 15 showing the sample hr and gene number.

The Livak method utilized a comparison between the test sample (keratinocytes co-culture samples) and a calibrator sample (keratinocyte growth media with MSSA476 control samples) using the 16S rRNA gene as a reference. For example, the 24 hr sample was the comparison between K-24, keratinocytes co-cultured with MSSA476 for 24 hrs (test sample C(T) value) and Km-24, keratinocyte growth media inoculated with MSSA476 for 24 hrs

(calibrator sample C(T) value) with reference genes C(T) value from both (K-24 and Km-24) to give the fold increase in expression of test gene.

Table 15 Calculations to determine fold increase in protein expression were performed using the Livak method.

Sample hr_Gene #	**6hr_12**	**6hr_13**	**24hr_13**
ΔCT Calibrator	9.73	9.73	9.73
ΔCT test	2.94	2.94	3.22
Normalize	-6.79	-6.79	-6.51
Fold increase	110.66	110.66	91.13

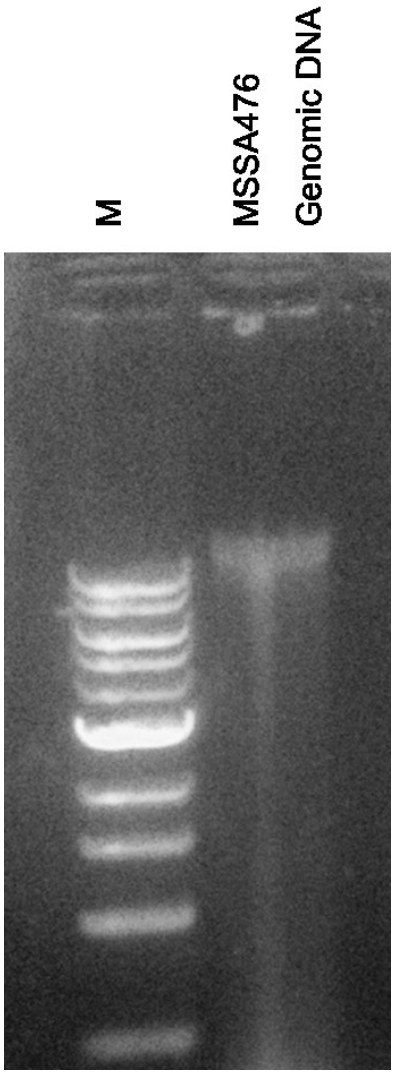


Figure 5 Genomic DNA.

Isolated genomic DNA on 1% agarose gel. Lane 1 (from left) is 1 Kb DNA marker (M) and lane 2 is 2 µl of isolated MSSA476 genomic DNA.

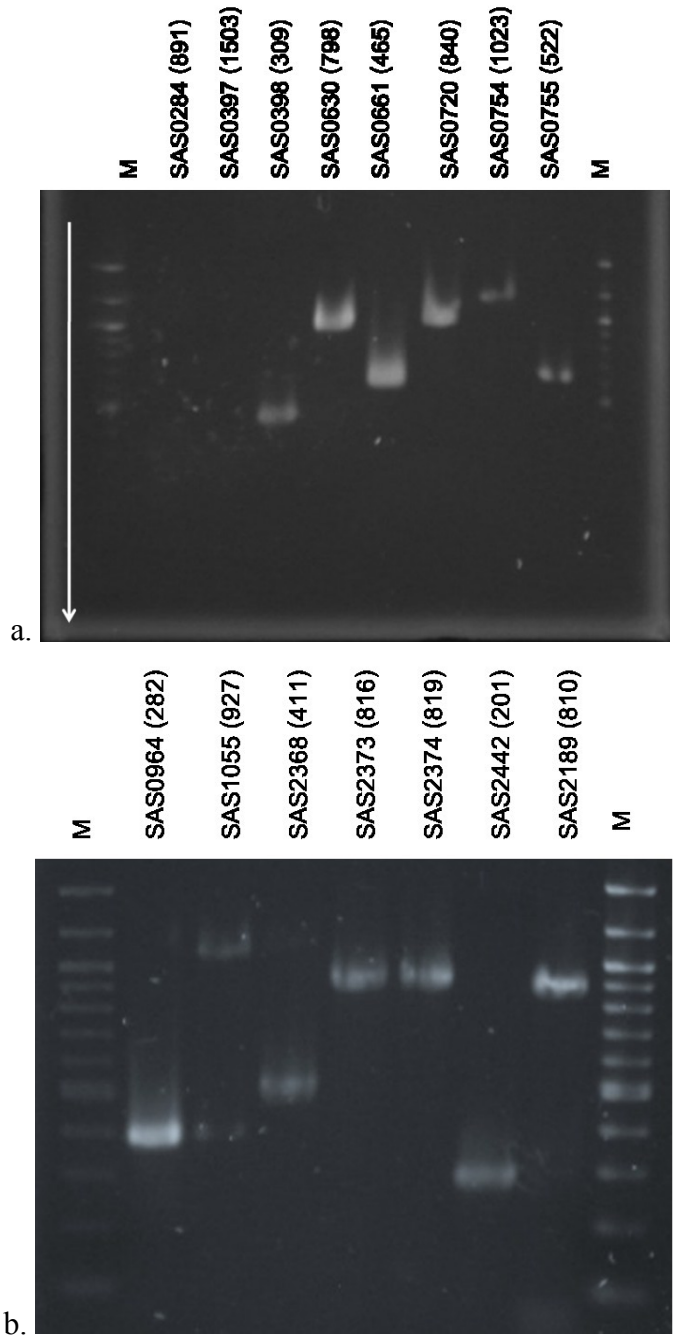
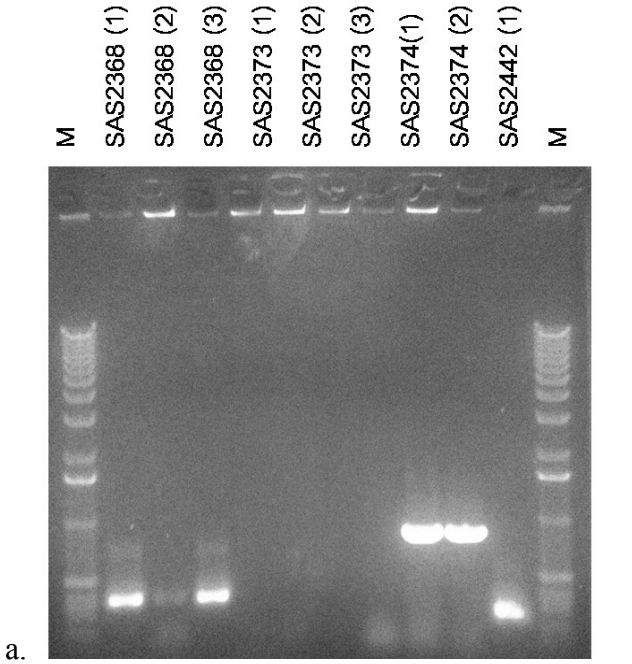
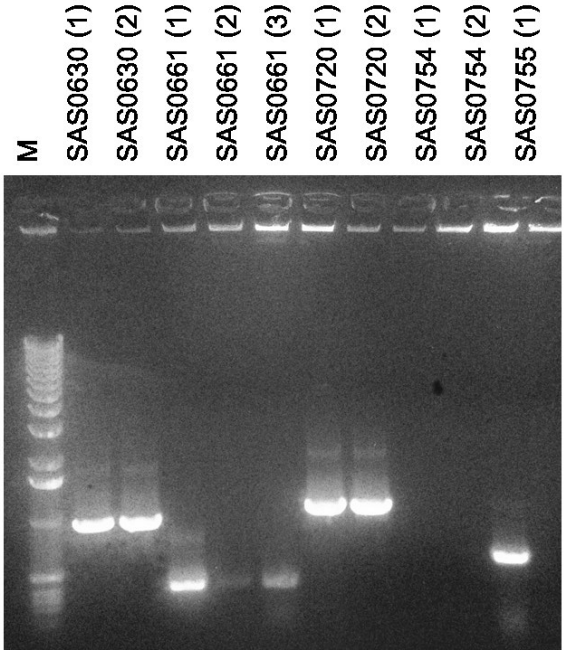


Figure 6 AttB PCR amplified product.

attB PCR product was detectable for 12 out of 14 genes electrophoresed on 1% agarose gel.



a.

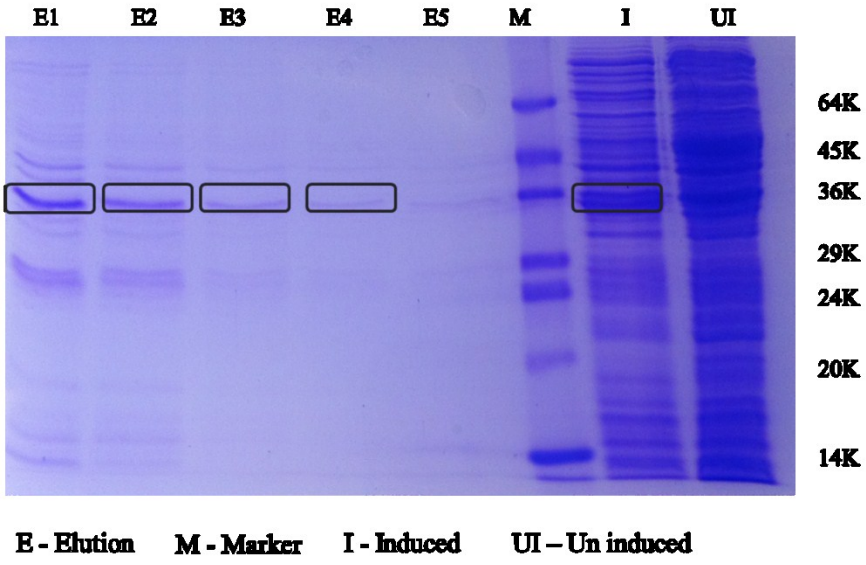


b.

Figure 7 Colony PCR amplification electrophoresed on 1% agarose gel.

Lane 1 (from left) is 1 Kb DNA marker; remaining lanes are colony PCR amplifications loaded on 1% agarose gel. Each lane was labeled with gene number of the clone followed by the colony number in closed brackets.

a.



b.

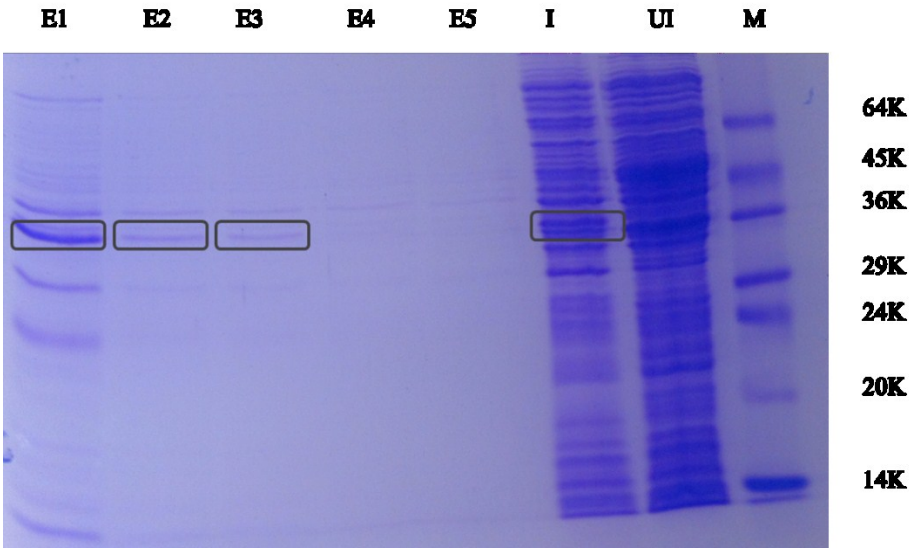


Figure 8 SDS gel showing protein bands.

Fig 8a shows an SDS gel with protein bands for gene SAS2373 and Fig 8b shows protein bands for gene SAS2374. Lane E1-5 are protein samples eluted with imidazole after purification; lanes labeled as I are induced Bl-21 clones with SAS2373 and SAS2374 genes, respectively; lanes labeled as UI are uninduced BL-21 clones with SAS2373 and SAS2374 genes, respectively; the lane labeled as M is protein marker.

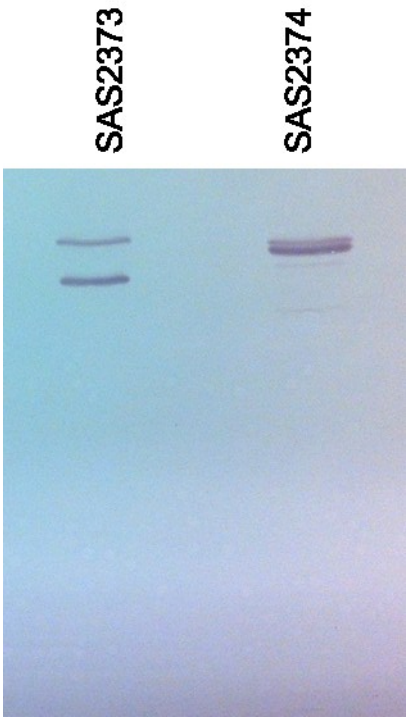


Figure 9 Western blot.

A nitrocellulose membrane exposed to substrate showed purple bands after 2 min. Lane 1 shows protein expression for gene SAS2373 and lane 2 shows protein expression for gene SAS2374.

a. 

b. 

Figure 10 *C. elegans* toxicity assay with BL-21 SAS2373 *E.coli* clone.

Petri plates with NGM media and BL-21 SAS2373 *E. coli* clone lawns were prepared. *C. elegans* was placed and allowed to grow for 4 days. The NGM plates containing IPTG and arabinose for induction of SAS2373 are shown in Fig 10a. The uninduced cells on plated without IPTG and arabinose are shown in Fig 10b. Both plates showed no bordering or reduction in number of *C. elegans*.

a. 

b. 

Figure 11 *C. elegans* toxicity assay with BL-21 SAS2374 *E.coli* clone.

Petri plates with NGM media and BL-21 SAS2374 *E. coli* clone lawns were inoculated with *C. elegans* and incubated for 4 days. The NGM plates containing IPTG and arabinose are shown in Fig 11a and uninduced cells (no IPTG or arabinose) are shown in Fig 11b. Neither plate showed bordering or reduction in number of *C. elegans*.

a. 

b. 

Figure 12 *C. elegans* toxicity assay with OP50 *E.coli* lawn.

Petri plates with NGM media and *E. coli* strain OP50 lawns were prepared as controls. *C. elegans* was placed and allowed to grow for 4 days. An NGM plate containing IPTG and arabinose is shown in Fig 12a and without IPTG and arabinose is shown in Fig 12b. Neither the induced nor the uninduced cultures showed bordering or reduction in number of *C. elegans*.

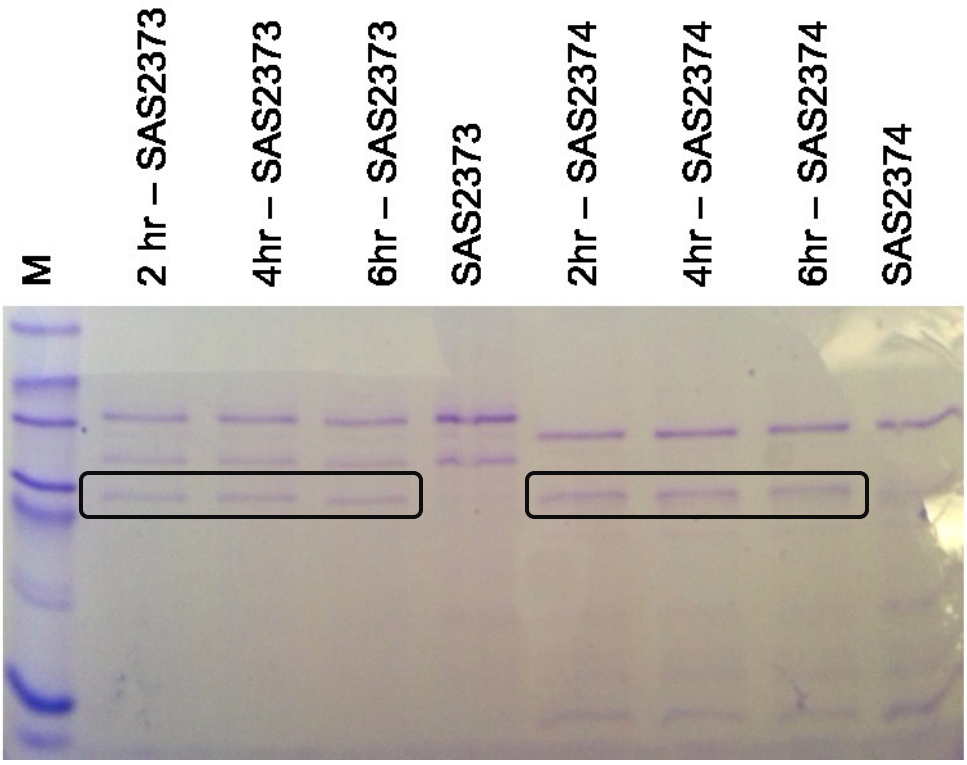


Figure 13 SDS gel loaded with AcTEV treated purified proteins.

AcTEV treated samples were loaded onto an SDS gel. Lane 1 (from left) is the size marker; lanes 2, 3 and 4 are SAS2373 protein sample treated with AcTEV protease for 2, 4, and 6 hrs, respectively. Lane 5 is SAS2373 protein sample without protease treatment. Lane 6, 7, and 8 are SAS2374 protein samples treated with AcTEC protease for 2, 4, and 6 hrs, respectively. Lane 9 is SAS2374 protein sample without protease treatment. All treated samples show an additional band around the 29 Kdal marker size (shown within black box) representing protease cleaved protein bands for both genes.

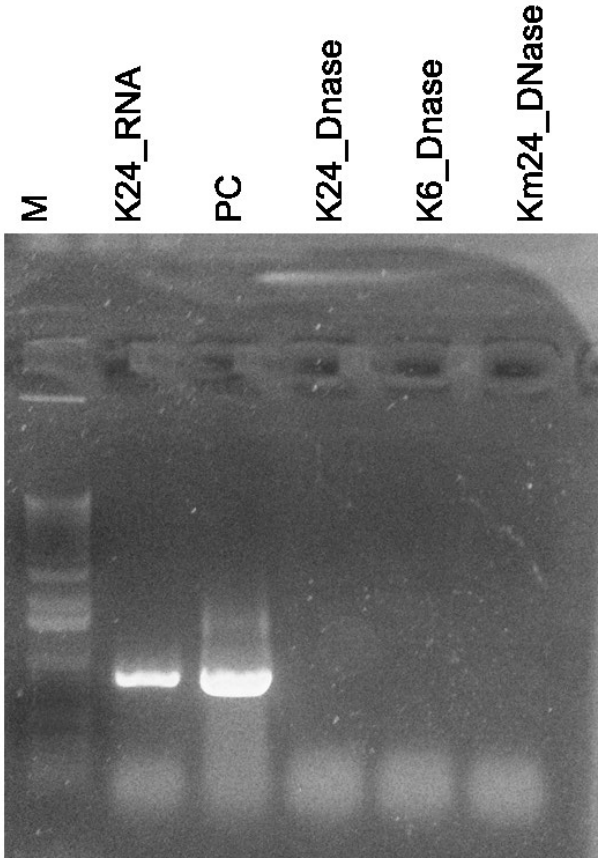
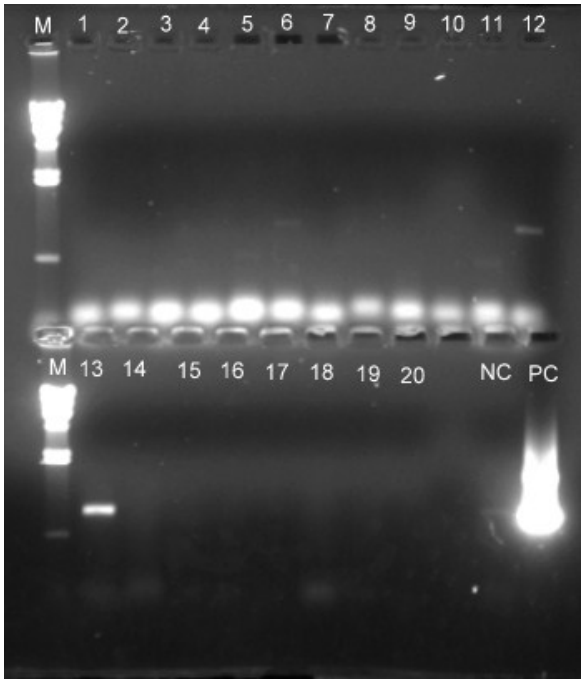


Figure 14 DNase-treated RNA samples.

DNase treatment: lane 1 (left most lane; M) is the size marker. The remaining lanes are samples after performing PCR with SAS2373 primers. Lane 2 shows the result of using the 24 hr keratinocyte RNA sample (K24_RNA) as the template for PCR. Lane 3 is the positive control (PC). Lane 4 is the 24 hr keratinocyte co-culture RNA sample treated with DNase (K24_DNase) and used as template for PCR. Lane 5 is the 6 hr keratinocyte co-culture RNA sample treated with DNase (K6_Dnase) and used as template for PCR. Lane 6 is the 24 hr keratinocyte (control) isolated RNA sample (Km24_DNase) and used as template for PCR.

a. 

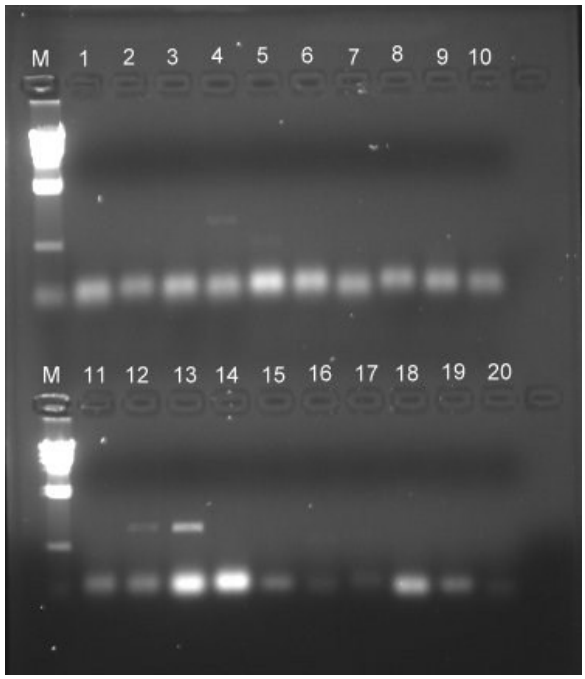
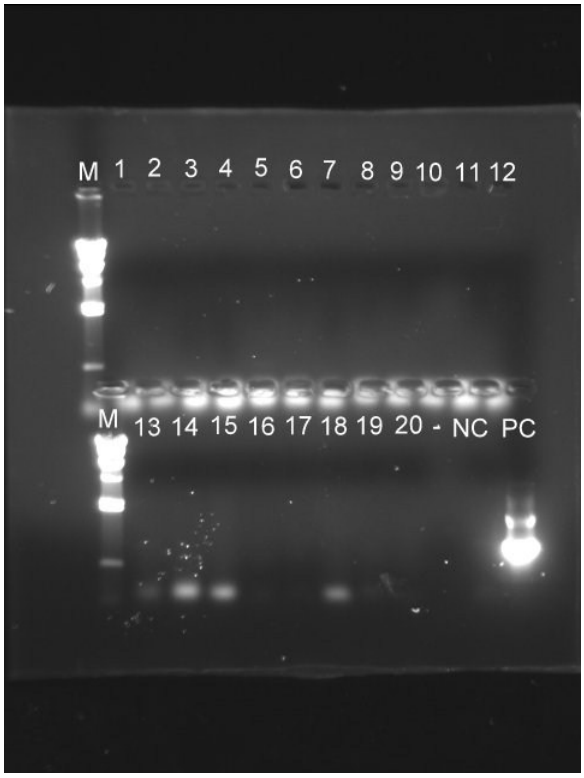
b. 

Figure 15 PCR amplification of 5 hr cDNA samples from MSSA476 co-cultured with keratinocytes.

Lane M is a 1 Kb DNA marker and the lane labeled PC is the positive control. NC is the negative control. cDNA samples of keratinocyte growth media inoculated with MSSA476 and incubated for 5 hours, Km-5 (Fig 15a) and keratinocytes co-cultured with MSSA476 for 5 hours, K-5 (Fig 15b) were used to amplify genes. DNA band sizes corresponded to the gene sizes and were identified with PCR amplification of all twenty samples listed in Table 10. Sample Km-5 showed positive amplification of gene numbers 5, 6, 9, 11, 12 &13. Sample K-5 showed positive amplification of gene number 4, 5, 12, 13, 16 and 17. Some of the gene amplifications resulted in very faint DNA bands corresponding to their respective gene size.

a. 

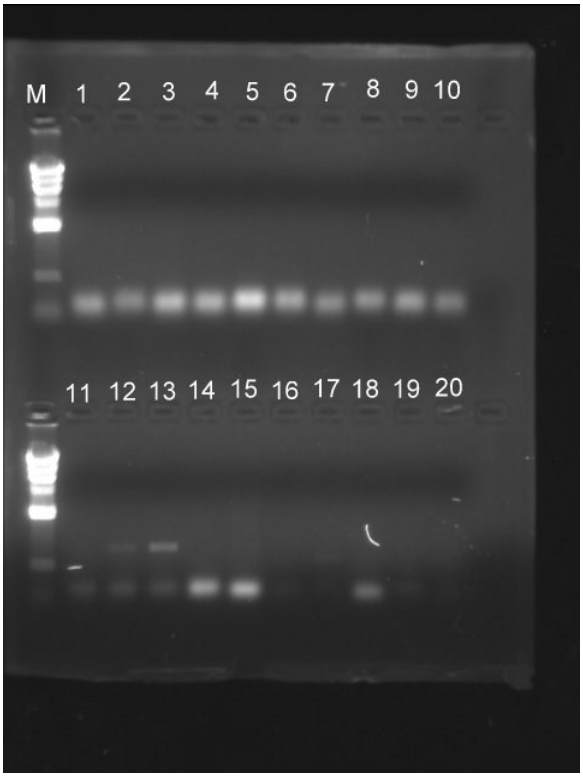
b. 

Figure 16 PCR amplification of 6 hrs cDNA samples from MSSA476 co-cultured with keratinocytes.

Lane M is 1 Kb DNA marker and the lane labeled as PC is positive control. NC is negative control. cDNA samples of keratinocyte growth media inoculated with MSSA476 and incubated for 6 hours, Km-6 (Fig 16a) and keratinocytes co-cultured with MSSA476 for 6 hrs, K-6 (Fig 16b) were used to amplify genes. DNA band sizes corresponding to the gene sizes were identified by PCR amplification of samples with the twenty genes listed in Table 10. Sample Km-6 was negative for all twenty gene amplifications. Sample K-6 showed positive amplification for gene number 12, 13 and 17. The product of gene 17 amplification was very faint.

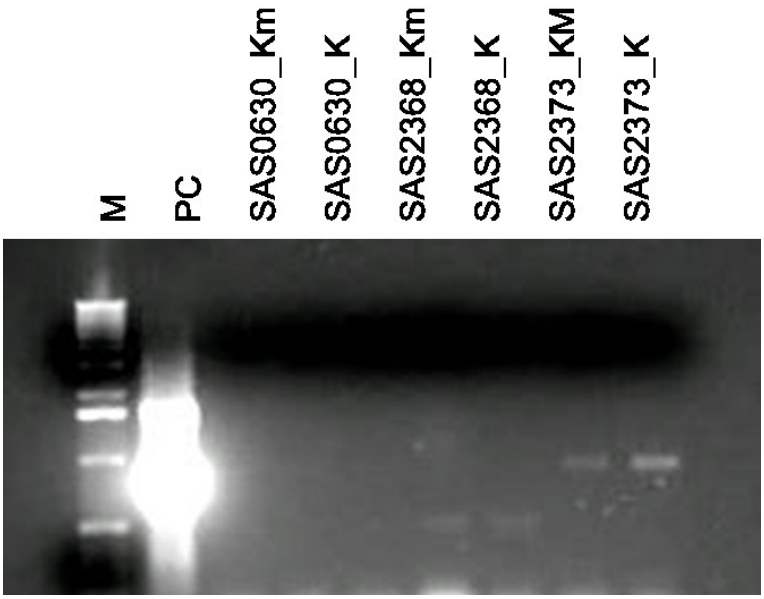
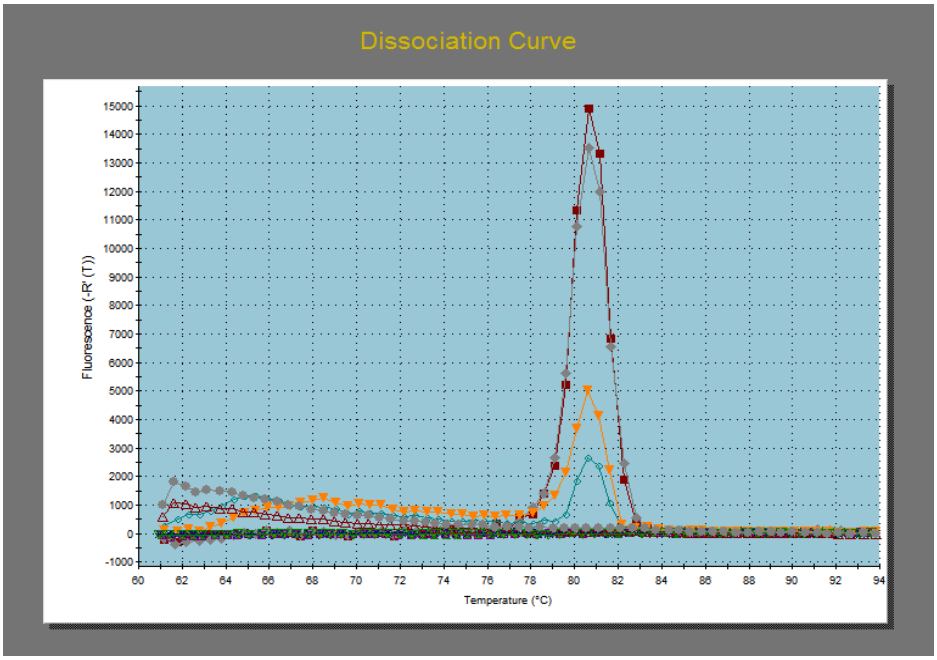


Figure 17 PCR amplification of the 24 hr cDNA samples from MSSA476 co-cultured with keratinocytes.

Lane 1 (far left) is the 1 Kb DNA marker; lane 2 is the positive control. cDNA samples of keratinocyte co-cultures with MSSA476 for 24 hrs (K24) and keratinocyte growth media inoculated with MSSA476 for 24 hrs (Km24) were used to amplify genes. DNA band sizes corresponding to gene sizes were identified in the PCR amplification products with SAS0630 (lane 3 & 4), SAS2368 (lane 5 & 6) and SAS2373 primers (lane 7 and 8).

a.



b.

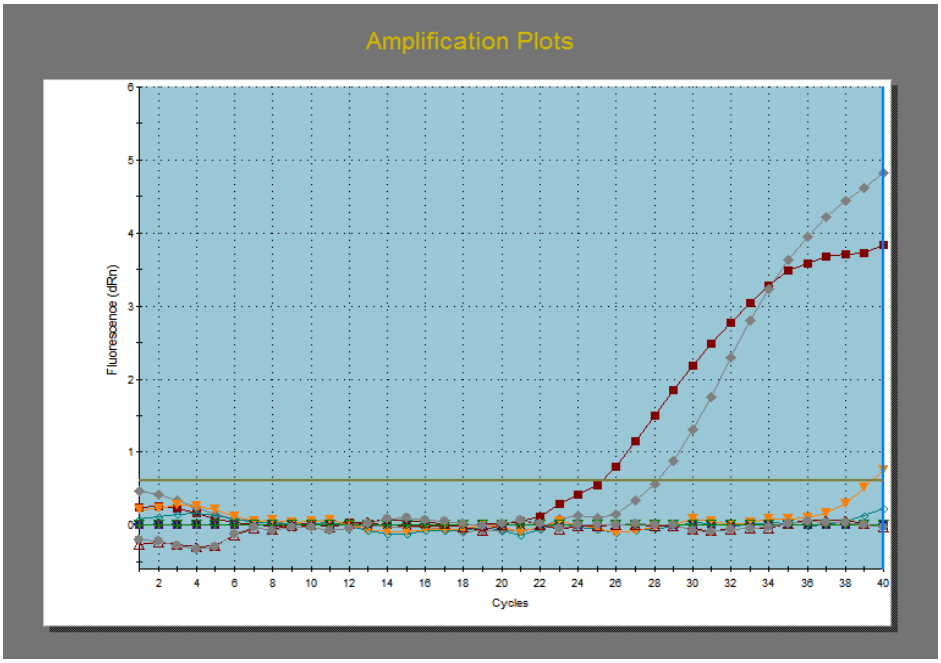
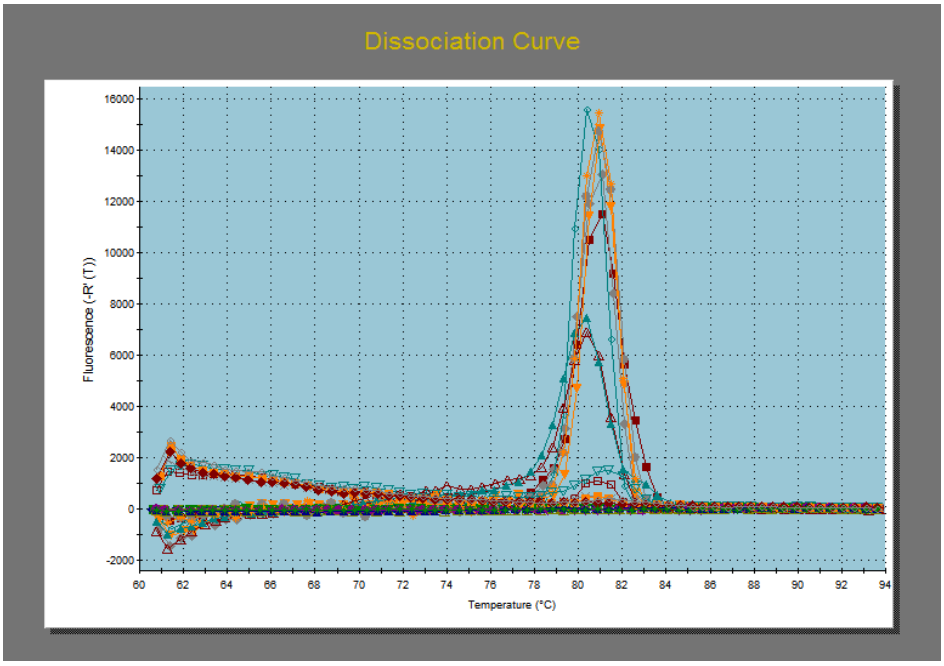


Figure 18 Dissociation curve and amplification plot of gene 2373.

Dissociation plots of gene 2373 showed amplification of the target gene. No background amplification could be detected. The amplification plot showed a positive amplification of gene SAS2373 (maroon and grey line) with K-24 cDNA sample and no amplification of SAS2373 (orange and green line) with Km-24 cDNA sample.

a.



b.

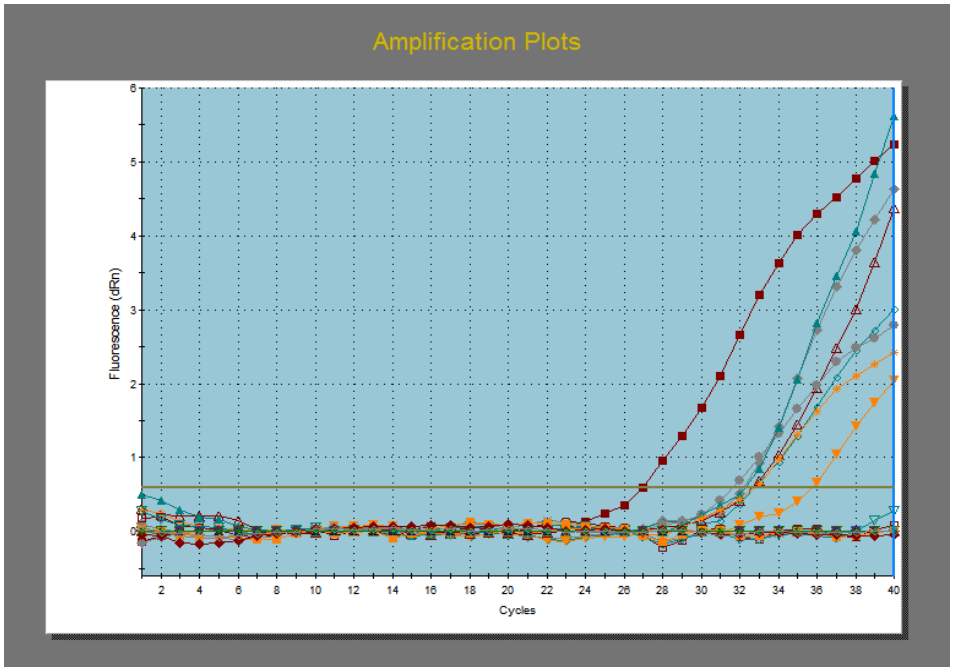


Figure 19 Dissociation curve and amplification plots for gene 2374.

Dissociation curve showed amplification of target genes without any background false positive amplifications. The amplification plot showed 8 lines corresponding to amplification of gene SAS2374 with K-24, Km-24, K-6 and Km-6 samples.

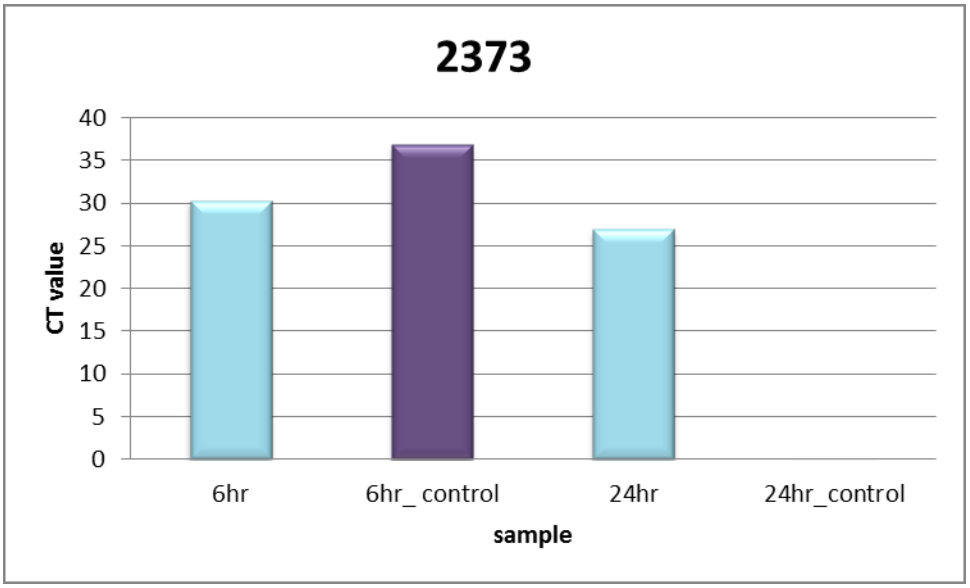


Figure 20 Graph showing C(T) values for gene sas2373 with 6 and 24 hr samples.

The graph shows the C(T) values for the gene *sas*2373 with K-6 and K-24 test samples (blue bars), and the control samples Km-6 and Km-24 (purple bars). The lower C(T) value is directly proportional to higher template availability, therefore, the lower C(T) value indicated higher concentration of template DNA. Km-24 showed no C(T) value indicating absence or very low concentration of template DNA and K-24 showed lowest C(T) value indicating highest concentration of DNA template among all sample.

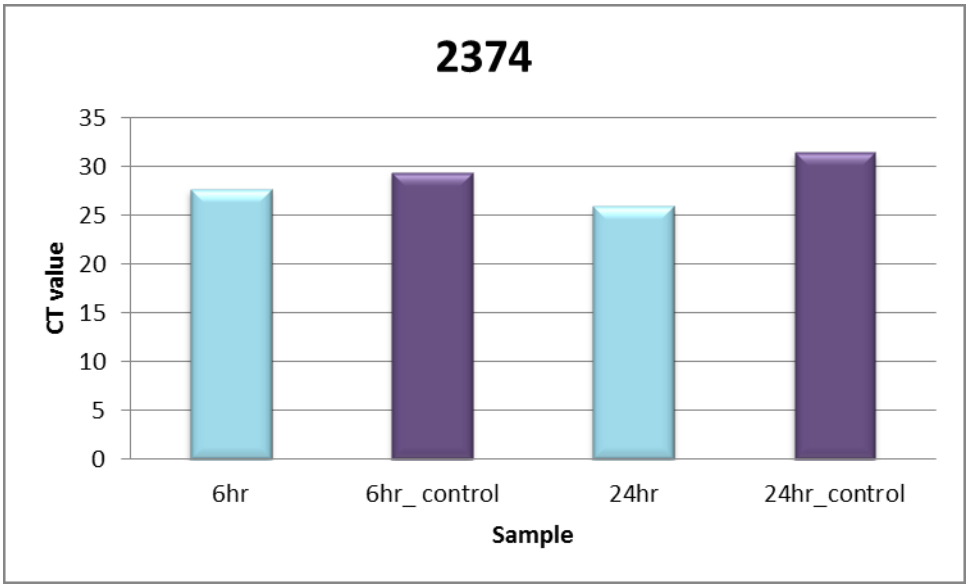


Figure 21 Graph showing C(T) value of gene sas2374 with 6 and 24hr samples.

The graph shows C(T) values for gene *sas*2374 with K-6 and K-24 test samples (blue bars) and Km-6 and Km-24 control samples (purple bars). The lower C(T) value is directly proportional to higher template availability; therefore, lower C(T) value indicated higher concentration of template DNA. K-24 showed lowest C(T) value of all indicating highest concentration of template DNA and Km-24 showed highest C(T) value indicating lowest amount of template DNA.

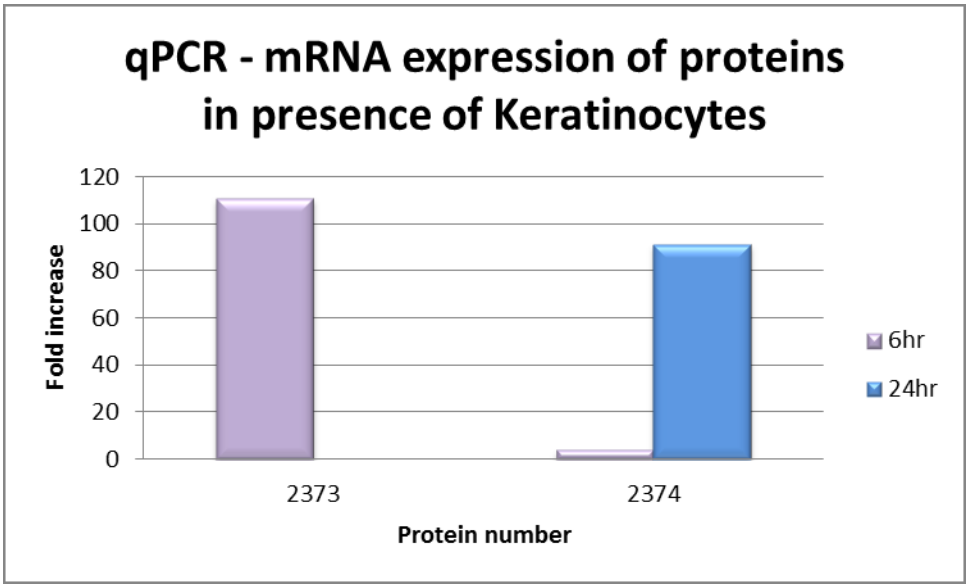


Figure 22 Fold increase in expression of mRNA during co-culture.

Using the Livak method, fold increase in expression of *sas*2373 and *sas*2374 mRNA are plotted.

Keratinocyte co-culture showed increased expression of both SAS2373 and SAS2374 mRNA.

SAS2374 showed 3.7-fold increased expression at 6 hrs and 91.13-fold increase at 24 hrs.

SAS2373 showed 110.66-fold increased expression at 6 hrs but no increase at 24 hrs.

CHAPTER 4

DISCUSSION

In this study, we investigated the involvement of twenty hypothetical exported proteins of MSSA476 in breaching the skin barrier and aiding the bacteria in invasion and causing deep tissue damage. Our study reveals two proteins with increased mRNA expression in presence of keratinocytes which may indicate their involvement in interacting with keratinocytes and aiding the bacterium to gain access to deeper tissue. These proteins may be relevant to MSSA476 in colonization, internalization, evading immune response, or in binding to TJs located in the epidermal layers aiding the bacteria in invasion.

Many different proteins have been shown to be involved in invasion of keratinocytes by *S. aureus* [16, 68]. Studies of adherence proteins have shown that adhering the pathogen to target cells aids infection of the host by facilitating colonization and invasion of host cells. For example, fibronectin-binding proteins (FnBPs) of *S. aureus*, isolated from human airway secretions, are involved in adherence of bacterium to undifferentiated or poorly differentiated epithelial cells [69]. In our study we observed increased expression of two proteins, SAS2373 and SAS2374, in the presence of keratinocytes. From protein BLAST, we observed 100% identity to CA-MRSA only for both proteins. Both genes are located adjacent to each other and the presence of another lipoprotein SAS2375 lies upstream of SAS2374, suggesting these protein to be tandem lipoproteins. SAS2375 is also observed to be unique to CA-MRSA, Further

computational analysis of these proteins included a search for the LIPOBOX in which we concluded that the proteins are not lipoproteins. Cumulatively, our observations suggest that these proteins may be associated with *S. aureus* interactions with keratinocytes.

Our study shows that SAS2373 mRNA is expressed in presence of keratinocytes and not in control sample at 24 hrs. This indicates that factors associated with the keratinocytes are triggering the expression of SAS2373. Following initiation of transcription, mRNA is translated and these proteins follow a specific pathway to be secreted from the cell [70]. For this process to take place, the bacteria may detect factors or signals associated with the host [71] and our results indicate factors or signals associated with keratinocytes have activated transcription of this gene.

SAS2373 expression at 6 hrs in co-culture showed expression of protein in control samples and samples with keratinocytes, but with a 111-fold greater increase in presence of keratinocytes. The increased expression of *sas*2373 at 6 hrs and the continued expression at 24 hrs suggested the up-regulation of transcription of genes was triggered by keratinocytes and not by media components.

SAS2374 gene expression showed 3.7-fold increase and 91.13-fold increase in expression in presence of keratinocytes at 6 and 24 hrs, respectively. The expression of gene SAS2374 was relatively low at 6 hrs compared to SAS2373 indicating the expression of SAS2374 may be position mediated by the co-expression of SAS2373 as described previously for *Saccharomyces cerevisiae* paired gene positioning within the ribosome biogenesis regulon [72]. A study showing the requirement of at least three high affinity fibronectin-binding protein repeats to successfully invade keratinocytes [68] and the involvement of two adjacent genes (fibronectin binding protein A & B) for adherence [69] stresses the involvement of more than a single protein in causing infection and interacting with host cells. A bacterium responding to its environment and with

changes to its environment, proteins must be up-regulated or down-regulated, depending on the environmental cues from its surroundings [72]. Expression of adjacent proteins in presence of keratinocytes suggests co-expression of these proteins which is supported by previous studies reporting co-expression of adjacent genes involved in similar functions [72].

SAS2373 and SAS2374 were not involved in antimicrobial activity as shown by microbial killing assay where no lysis of normal flora was observed and therefore, suggesting these proteins are not involved in inhibiting normal flora for colonization.

Our results with *C. elegans* indicated no behavioral changes in the presence of SAS2373 or SAS2374. *C. elegans* are capable of showing various behavioral changes when harmful or toxic chemicals are in its environment [57]. Noticing no change in the behavior pattern and no reduction in the number of animals in test plate suggested these proteins were not identified by the nematode as harmful to the adult animal or its eggs. As previously shown [26] in this laboratory, one hypothetical exported protein from MSSA476 had inhibitory effects on *C. elegans*. It caused reduction in number and bordering with *C. elegans* exposure to induced protein. Both of these characteristics were absent from the current study.

From the microbial killing assay, *C. elegans* studies, and mRNA expression results, it was observed that the expressed proteins are directly linked to keratinocytes. Many other proteins described in previous studies have shown a role in colonization by the production of FnBPs which not only help in invasion of the bacteria with high affinity FnBP repeats, but also without these repeats or at low affinity. These proteins can help in adhering, followed by colonization with undifferentiated epithelial cells [69], suggesting that non-toxic proteins may be involved in virulence by other mechanisms such as adherence to the target cells and colonization. Colonization is a strategic move to ensure that a bacterium is available to cause infection when

conditions are favorable [68]. In our study, we observed that keratinocytes cells were not rounded or otherwise visibly affected by MSSA476 co-culture, suggesting that the secreted proteins are not cytotoxic or in any other way disrupt the TJs of the keratinocytes. However, our observation may be consistent with adhesion to the keratinocytes. Samples collected after washing the cells with PBS (therefore, most of the free floating bacterial cells were removed and only adhered bacterial cells remained), did not show rounding as they would if TJs were broken or indicating any signs of invasion of the keratinocytes. It was obvious, however, that the bacteria were bound to the keratinocytes.

Keratinocytes not only differentiate to form the epidermal outer protective layer but they are capable of producing antimicrobial peptides, cytokines, and the junction between the cells makes the epidermal layer more rigid and prevents microbes from freely entering our bodies. The immune responses associated with keratinocytes are well documented both in in vivo and in vitro conditions.

TJs in the epidermal layers, mostly located in the stratum granulosum, act as a barrier and help retain water [73]. Many modes of bacterial invasion or interaction with keratinocytes have been described. Some bacterial pathogens are capable of adhering to the junctions and internalizing while other pathogens can directly bind to proteins associated with TJ and destroy them to gain access to deeper tissues [74]. In the current study, we observed that keratinocytes were not visibly affected in presence of MSSA476 for 24 hrs under the conditions used suggesting the proteins in question may not be interacting with TJs. If either protein was associated with binding to the TJ for the purpose of invading the cell, the keratinocytes should have been affected by 24 hrs. If either of the proteins was capable of specifically destroying one or more of the TJ proteins, we would have observed keratinocytes cell rounding or deformation

due to destruction of TJs between neighboring cells. These observations not only suggest that these proteins are not involved in binding to TJs but also that MSSA476 in general was unable to interact or destroy TJs located on keratinocytes by any means under these conditions.

An important observation that was made was in areas of the keratinocyte monolayer where the cells did not grow sufficiently to complete the monolayer. The effect of MSSA476 on small colonies of cells (3 cell colonies) was that the cells showed rounding or deformation indicating proteins associated with the MSSA476 may affect wounded areas where the TJs or keratinocyte cells are separated or open to bacterial access.

Keratinocytes in cell culture can differentiate to produce antimicrobial peptides by increasing the amount of calcium (1.15 mM) in medium [43]. With this increase in calcium, keratinocytes were shown in previous studies to express increased amounts of antimicrobial peptides after 24 hrs of co-incubation intracellular *S. aureus* [43]. Even with low amounts (0.15mM) of calcium chloride in solution, antimicrobial peptides were expressed at low levels constantly for 24 hrs [43]. Keratinocyte growth media purchased from Zen-Bio has very low concentrations of calcium chloride (approximately 0.016 mM) and may not have enabled keratinocytes to differentiate to produce high amounts of antimicrobial peptides. This observation suggests that SAS2373 and sas2374 were not expressed in response to antimicrobial peptides.

RNase 7 is another type of antimicrobial produced by keratinocytes, primarily in the stratum corneum layers. Studies have shown broad spectrum antimicrobial activity in vitro with 2-8 μM RNase 7 concentration for 3 hrs towards various bacterial strains [49]. The in vitro study concentration is high when compared to an in vivo study measuring RNase 7 from different area of the body to be between 0.3-3.5 ng/cm^2, depending upon individual and site tested on the body

[48]. These studies suggest that the amount of RNase 7 produced by our body may not be as high as was shown to be necessary for antimicrobial activity towards various bacterial cultures. The fact that the RNase 7 is mostly found in the stratum corneum may also suggest the requirement for a break in the skin for *Staphylococcus aureus* to gain access to host tissues. In culture, high amounts of RNase 7 have been detected in the supernatant after 16 hrs. Rnase 7 should degrade any RNA from the sample or media. The fact that we identified mRNA from keratinocyte co-culture samples suggests that the RNA was not degraded by RNase 7. It is possible that the initial step to stabilize RNA samples using phenol/ethanol mixture may have inhibited RNase 7 activity.

Keratinocytes produce cytokines at low levels all the time but in case of a breech the cytokine production is elevated and triggers the immune response. In culture, primary keratinocytes are known to produce IL-1, IL-6, and TNF-α. They can be induced by toxins, skin injuries, UV light, bacterial cell wall components, etc. Studies growing keratinocytes in planktonic culture media (PCM) and biofilm culture media (BCM) showed high cytokine production with PCM compared to BCM [75]. IL-1 production was not affected by varying amounts of calcium chloride in solution and was produced by both differentiated and undifferentiated cells [76]. IL-1 and TNF-α played a role in initiating a cytokine cascade which would trigger the immune response not only to fight off invading microbes but also to conduct repairs caused to the skin barrier. In our experiments, the presence of MSSA476 may have triggered the production of cytokines which in turn may have caused expression of SAS2373 and SAS2374. Many bacterial proteases are known to bind to TNF-α and deactivate it. Some bacterial proteins and toxins are involved in binding to IL-1 and utilize cytokine derivatives as growth factors [76]. IL-1, TNF-α, and IL-6 are capable of inducing production of antimicrobial

peptides in keratinocytes. Though low amounts of calcium may not initiate production of antimicrobial peptides, the cytokines released in response to the bacteria may induce the production of antimicrobial peptides [44].

Therefore, more investigation is required to understand the relationship between the expression of SAS2373 and SAS2374 proteins in the presence of keratinocytes and the role they play. Increased expression may be in response to antimicrobial peptides or cytokines produced by keratinocytes, or may be involved in binding to exposed keratinocytes at the site of a wound.

Future directions may include growing keratinocytes and causing a wound on the cultured cells via a pipette tip, inoculating MSSA476, and observing expression at the mRNA level. ELISA tests could be performed on co-cultured samples to test for expression of cytokines. An antimicrobial peptides gene expression study could be performed with keratinocytes in presence of MSSA476.

REFERENCES

1. Kluytmans J, van Belkum A, Verbrugh H: **Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks.** *Clinical microbiology reviews* 1997, **10:**505-520.
2. Gordon RJ, Lowy FD: **Pathogenesis of methicillin-resistant Staphylococcus *aureus* infection.** *Clin Infect Dis* 2008, **46 Suppl 5:**S350-359.
3. Malachowa N, Kobayashi SD, DeLeo FR: **Community-associated methicillin-resistant *Staphylococcus aureus* and athletes.** *The Physician and sportsmedicine* 2012, **40:**13-21.
4. Ladhani S, Joannou CL, Lochrie DP, Evans RW, Poston SM: **Clinical, microbial, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded-skin syndrome.** *Clinical microbiology reviews* 1999, **12:**224-242.
5. Biedenbach DJ, Moet GJ, Jones RN: **Occurrence and antimicrobial resistance pattern comparisons among bloodstream infection isolates from the SENTRY Antimicrobial Surveillance Program (1997-2002).** *Diagnostic microbiology and infectious disease* 2004, **50:**59-69.
6. Dinges MM, Orwin PM, Schlievert PM: **Exotoxins of Staphylococcus *aureus*.** *Clinical microbiology reviews* 2000, **13:**16-34, table of contents.
7. McBryde ES, Bradley LC, Whitby M, McElwain DL: **An investigation of contact transmission of methicillin-resistant *Staphylococcus aureus*.** *The Journal of hospital infection* 2004, **58:**104-108.
8. Carr JH: **Under a very high magnification of 20,000x, this scanning electron micrograph (SEM) shows a strain of *Staphylococcus aureus* bacteria taken from a vancomycin intermediate resistant culture (VISA.** In *Book Under a very high magnification of 20,000x, this scanning electron micrograph (SEM) shows a strain of Staphylococcus aureus bacteria taken from a vancomycin intermediate resistant culture (VISA* (Editor ed.^eds.). City: CDC; 2001.
9. Moroney SM, Heller LC, Arbuckle J, Talavera M, Widen RH: **Staphylococcal cassette chromosome mec and Panton-Valentine leukocidin characterization of methicillin-resistant *Staphylococcus aureus* clones.** *Journal of clinical microbiology* 2007, **45:**1019-1021.
10. Loomba PS, Taneja J, Mishra B: **Methicillin and Vancomycin Resistant *S. aureus* in Hospitalized Patients.** *Journal of global infectious diseases* 2010, **2:**275-283.
11. Sydnor ER, Perl TM: **Hospital epidemiology and infection control in acute-care settings.** *Clinical microbiology reviews* 2011, **24:**141-173.
12. Waness A: **Revisiting Methicillin-Resistant *Staphylococcus aureus* Infections.** *Journal of global infectious diseases* 2010, **2:**49-56.

13. Kluytmans JA, Mouton JW, Ijzerman EP, Vandenbroucke-Grauls CM, Maat AW, Wagenvoort JH, Verbrugh HA: **Nasal carriage of Staphylococcus *aureus* as a major risk factor for wound infections after cardiac surgery.** *The Journal of infectious diseases* 1995, **171:**216-219.
14. **MRSA Statistics**
15. Zetola N, Francis JS, Nuermberger EL, Bishai WR: **Community-acquired meticillin-resistant *Staphylococcus aureus*: an emerging threat.** *Lancet Infect Dis* 2005, **5:**275-286.
16. David MZ, Daum RS: **Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic.** *Clinical microbiology reviews* 2010, **23:**616-687.
17. Perichon B, Courvalin P: **VanA-type vancomycin-resistant *Staphylococcus aureus*.** *Antimicrobial agents and chemotherapy* 2009, **53:**4580-4587.
18. Cui L, Murakami H, Kuwahara-Arai K, Hanaki H, Hiramatsu K: **Contribution of a thickened cell wall and its glutamine nonamidated component to the vancomycin resistance expressed by *Staphylococcus aureus* Mu50.** *Antimicrobial agents and chemotherapy* 2000, **44:**2276-2285.
19. Stacey AR, Endersby KE, Chan PC, Marples RR: **An outbreak of methicillin resistant *Staphylococcus aureus* infection in a rugby football team.** *Br J Sports Med* 1998, **32:**153-154.
20. Lowy FD, Aiello AE, Bhat M, Johnson-Lawrence VD, Lee MH, Burrell E, Wright LN, Vasquez G, Larson EL: ***Staphylococcus aureus* colonization and infection in New York State prisons.** *The Journal of infectious diseases* 2007, **196:**911-918.
21. Gillet Y, Issartel B, Vanhems P, Fournet JC, Lina G, Bes M, Vandenesch F, Piemont Y, Brousse N, Floret D, Etienne J: **Association between Staphylococcus *aureus* strains carrying gene for Panton-Valentine leukocidin and highly lethal necrotising pneumonia in young immunocompetent patients.** *Lancet* 2002, **359:**753-759.
22. Vandenesch F, Naimi T, Enright MC, Lina G, Nimmo GR, Heffernan H, Liassine N, Bes M, Greenland T, Reverdy ME, Etienne J: **Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence.** *Emerg Infect Dis* 2003, **9:**978-984.
23. Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, et al: **Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101:**9786-9791.
24. Luong TT, Ouyang S, Bush K, Lee CY: **Type 1 capsule genes of *Staphylococcus aureus* are carried in a staphylococcal cassette chromosome genetic element.** *Journal of bacteriology* 2002, **184:**3623-3629.
25. Katayama Y, Takeuchi F, Ito T, Ma XX, Ui-Mizutani Y, Kobayashi I, Hiramatsu K: **Identification in methicillin-susceptible *Staphylococcus* hominis of an active primordial mobile genetic element for the staphylococcal cassette chromosome mec of methicillin-resistant *Staphylococcus aureus*.** *Journal of bacteriology* 2003, **185:**2711-2722.

26. Kumar DKV, Dannelly DHK: **Cloning and characterization of SAS1738, a hypothetical exported protein from community associated strain of *Staphylococcus aureus*.** In *Book Cloning and characterization of SAS1738, a hypothetical exported protein from community associated strain of Staphylococcus aureus* (Editor ed.^eds.). pp. 91. City: Indiana State Univeristy; 2010:91.
27. Wickett RR, Visscher MO: **Structure and function of the epidermal barrier.** *American Journal of Infection Control* 2006, **Vol. 34:** S98-S110.
28. Salyers AA, Wilson BA, Whitt DD: *Bacterial Pathogenesis : A Molecular Approach (3rd Edition).* Washington, DC, USA: ASM Press; 2011.
29. Del Rosso JQ, Levin J: **The clinical relevance of maintaining the functional integrity of the stratum corneum in both healthy and disease-affected skin.** *The Journal of clinical and aesthetic dermatology* 2011, **4:**22-42.
30. *Textbook of Dermatology.* Blackwell Publishing Ltd; 2010.
31. Bensouilah J, Buck P: *Aromadermatology - Aromatherapy in the treatment and care of common skin conditions.* Radcliffe Publishing; 2006.
32. Elis PM, Feingold KR: *Skin Barrier.* New York: Taylor & Francis group; 2006.
33. Samantha Lin, Bingcheng Wang, Getsios S: **Eph/ephrin signaling in epidermal differentiation and disease.** *Seminars in Cell & Developmental Biology* 2012, **23:**92–101.
34. Kisich KO, Howell MD, Boguniewicz M, Heizer HR, Watson NU, Leung DY: **The constitutive capacity of human keratinocytes to kill *Staphylococcus aureus* is dependent on beta-defensin 3.** *The Journal of investigative dermatology* 2007, **127:**2368-2380.
35. Esche C, de Benedetto A, Beck LA: **Keratinocytes in atopic dermatitis: inflammatory signals.** *Current allergy and asthma reports* 2004, **4:**276-284.
36. Ansel J, Perry P, Brown J, Damm D, Phan T, Hart C, Luger T, Hefeneider S: **Cytokine modulation of keratinocyte cytokines.** *The Journal of investigative dermatology* 1990, **94:**101S-107S.
37. Mempel M, Voelcker V, Kollisch G, Plank C, Rad R, Gerhard M, Schnopp C, Fraunberger P, Walli AK, Ring J, et al: **Toll-like receptor expression in human keratinocytes: nuclear factor kappaB controlled gene activation by Staphylococcus *aureus* is toll-like receptor 2 but not toll-like receptor 4 or platelet activating factor receptor dependent.** *The Journal of investigative dermatology* 2003, **121:**1389-1396.
38. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: *Molecular Biology of the Cell, Fourth Edition.* New York: Garland Science; 2002.
39. O'Neill CA, Garrod D: **Tight junction proteins and the epidermis.** *Experimental dermatology* 2011, **20:**88-91.
40. Pummi K, Malminen M, Aho H, Karvonen SL, Peltonen J, Peltonen S: **Epidermal tight junctions: ZO-1 and occludin are expressed in mature, developing, and affected skin and in vitro differentiating keratinocytes.** *The Journal of investigative dermatology* 2001, **117:**1050-1058.
41. Brandner JM, Haftek M, Niessen CM: **Adherens Junctions, Desmosomes and Tight Junctions in Epidermal Barrier Function.** *The Open Dermatology Journal* 2010, **4:**14-20.

42. Young P, Boussadia O, Halfter H, Grose R, Berger P, Leone DP, Robenek H, Charnay P, Kemler R, Suter U: **E-cadherin controls adherens junctions in the epidermis and the renewal of hair follicles.** *The EMBO journal* 2003, **22:**5723-5733.
43. Menzies BE, Kenoyer A: ***Staphylococcus aureus* infection of epidermal keratinocytes promotes expression of innate antimicrobial peptides.** *Infection and immunity* 2005, **73:**5241-5244.
44. Andor P, Istvan N, Lajos K: **Innate Immunity in the Skin: How Keratinocytes Fight Against Pathogens.** *Current Immunology Reviews* 2005, **1:**29-42.
45. Yeung A, Gellatly S, Hancock R: **Multifunctional cationic host defence peptides and their clinical applications.** *Cellular and Molecular Life Sciences* 2011, **68:**2161-2176.
46. Prenner EJ, Lewis RN, Jelokhani-Niaraki M, Hodges RS, McElhaney RN: **Cholesterol attenuates the interaction of the antimicrobial peptide gramicidin S with phospholipid bilayer membranes.** *Biochimica et biophysica acta* 2001, **1510:**83-92.
47. Pazgier M, Hoover DM, Yang D, Lu W, Lubkowski J: **Human beta-defensins.** *Cellular and molecular life sciences : CMLS* 2006, **63:**1294-1313.
48. Koten B, Simanski M, Glaser R, Podschun R, Schroder JM, Harder J: **RNase 7 contributes to the cutaneous defense against Enterococcus faecium.** *PloS one* 2009, **4:**e6424.
49. Harder J, Schroder JM: **RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin.** *The Journal of biological chemistry* 2002, **277:**46779-46784.
50. Groisman EA: *principles of bacterial pathogensesis.* San Diego, Callifornia: Academic Press; 2001.
51. Lee CA, Jones BD, Falkow S: **Identification of a Salmonella typhimurium invasion locus by selection for hyperinvasive mutants.** *Proceedings of the National Academy of Sciences of the United States of America* 1992, **89:**1847-1851.
52. Hromockyj AE, Maurelli AT: **Identification of Shigella invasion genes by isolation of temperature-regulated inv::lacZ operon fusions.** *Infection and immunity* 1989, **57:**2963-2970.
53. van Roosmalen ML, Geukens N, Jongbloed JD, Tjalsma H, Dubois JY, Bron S, van Dijl JM, Anne J: **Type I signal peptidases of Gram-positive bacteria.** *Biochimica et biophysica acta* 2004, **1694:**279-297.
54. Hornef MW, Wick MJ, Rhen M, Normark S: **Bacterial strategies for overcoming host innate and adaptive immune responses.** *Nature immunology* 2002, **3:**1033-1040.
55. Maresso AW, Garufi G, Schneewind O: **Bacillus anthracis secretes proteins that mediate heme acquisition from hemoglobin.** *PLoS Pathog* 2008, **4:**e1000132.
56. Wood WB: *The nematode Caenorhabditis elegans.* New York, NY: Cold Spring Harbor Laboratory Press.; 1988.
57. Rankin CH: **From gene to identified neuron to behaviour in *Caenorhabditis*.** *Nature reviews Genetics* 2002, **3:**622-630.
58. Kall L, Krogh A, Sonnhammer EL: **A combined transmembrane topology and signal peptide prediction method.** *Journal of molecular biology* 2004, **338:**1027-1036.
59. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Eng* 1997, **10:**1-6.

60. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *Journal of molecular biology* 2004, **340:**783-795.
61. Viklund H, Bernsel A, Skwark M, Elofsson A: **SPOCTOPUS: a combined predictor of signal peptides and membrane protein topology.** *Bioinformatics* 2008, **24:**2928-2929.
62. Sonnhammer EL, von Heijne G, Krogh A: **A hidden Markov model for predicting transmembrane helices in protein sequences.** *Proc Int Conf Intell Syst Mol Biol* 1998, **6:**175-182.
63. Weisberg RA, Enquist LW, Foeller C, Landy A: **Role for DNA homology in site-specific recombination. The isolation and characterization of a site affinity mutant of coliphage lambda.** *Journal of molecular biology* 1983, **170:**319-342.
64. Tokuda H, Matsuyama S: **Sorting of lipoproteins to the outer membrane in *E. coli*.** *Biochimica et biophysica acta* 2004, **1693:**5-13.
65. Jahn CE, Charkowski AO, Willis DK: **Evaluation of isolation methods and RNA integrity for bacterial RNA quantitation.** *Journal of microbiological methods* 2008, **75:**318-324.
66. Franca A, Melo LD, Cerca N: **Comparison of RNA extraction methods from biofilm samples of *Staphylococcus* epidermidis.** *BMC Res Notes* 2011, **4:**572.
67. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.** *Methods* 2001, **25:**402-408.
68. Edwards AM, Potter U, Meenan NA, Potts JR, Massey RC: ***Staphylococcus aureus* keratinocyte invasion is dependent upon multiple high-affinity fibronectin-binding repeats within FnBPA.** *PloS one* 2011, **6:**e18899.
69. Mongodin E, Bajolet O, Cutrona J, Bonnet N, Dupuit F, Puchelle E, de Bentzmann S: **Fibronectin-binding proteins of *Staphylococcus* aureus are involved in adherence to human airway epithelium.** *Infection and immunity* 2002, **70:**620-630.
70. Schneewind O, Missiakas DM: **Protein secretion and surface display in Gram-positive bacteria.** *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2012, **367:**1123-1139.
71. Salis H, Tamsir A, Voigt C: **Engineering bacterial signals and sensors.** *Contributions to microbiology* 2009, **16:**194-225.
72. Arnone JT, McAlear MA: **Adjacent gene pairing plays a role in the coordinated expression of ribosome biogenesis genes MPP10 and YJR003C in Saccharomyces cerevisiae.** *Eukaryotic cell* 2011, **10:**43-53.
73. Rickelt S, Franke WW, Doerflinger Y, Goerdt S, Brandner JM, Peitsch WK: **Subtypes of melanocytes and melanoma cells distinguished by their intercellular contacts: heterotypic adherens junctions, adhesive associations, and dispersed desmoglein 2 glycoproteins.** *Cell and tissue research* 2008, **334:**401-422.
74. Guttman JA, Finlay BB: **Tight junctions as targets of infectious agents.** *Biochimica et biophysica acta* 2009, **1788:**832-841.
75. Secor PR, James GA, Fleckman P, Olerud JE, McInnerney K, Stewart PS: ***Staphylococcus aureus* Biofilm and Planktonic cultures differentially impact gene expression, mapk phosphorylation, and cytokine production in human keratinocytes.** *BMC microbiology* 2011, **11:**143.
76. Henderson B, Oyston P: *Bacterial Evasion of Host Immune Responses.* West Nyack, NY, USA: Cambridge University Press; 2003.

77. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *Journal of molecular biology* 2001, **305:**567-580.

APPENDIX A: COMPUTATIONAL ANALYSIS

Gene SAS1678 was identified on geneDB website and according to SignalP 2.0 version SAS1678 is exported proteins. SAS1678 was cloned and expressed using gateway cloning technique. SDS gel (fig A1) showed bands corresponding to SAS1678 protein size and bands were detected on western blot as well.

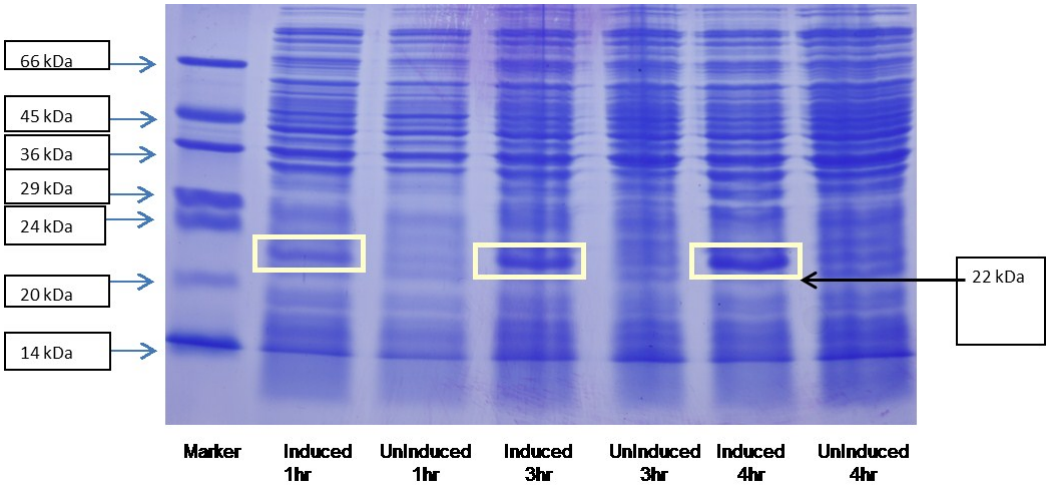


Figure A 1 SDS gel shows protein bands for induced and uninduced SAS1678 BL-21clone. Induced and uninduced SAS1678 BL-21 clone pellet dissolved in sample buffer and loaded onto the SDS gel. White rectangle shows highlighted 22 KDa protein corresponding to SAS1678 protein size.

Computational analysis of SAS1678 with Signal P 3 version (fig A2) and Phobius (fig A3) showed transmembrane domain, no signal peptide, and cytoplasmic region which suggested this protein is non – secretory and stays inside the cell, in the cytoplasm. These results showed a

need to further investigate secreted proteins from MSSA476 using various computational software's.

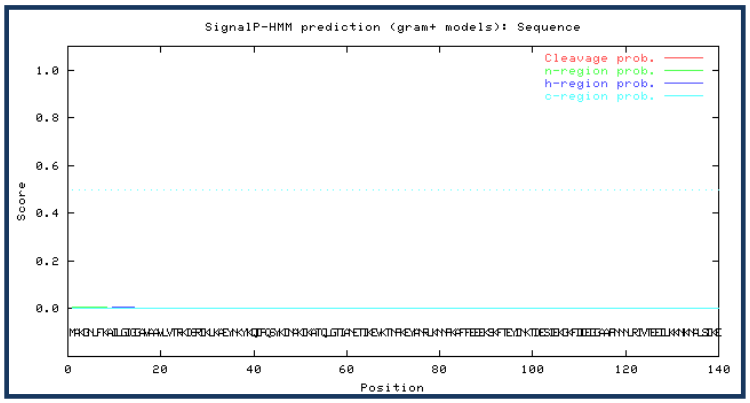


Figure A 2 Signal P 3.0 version plot for SAS1678.

Computational analysis of SAS1678 with Signal P.3 version software indicating test protein probability of signal peptide or cleavage site is zero.

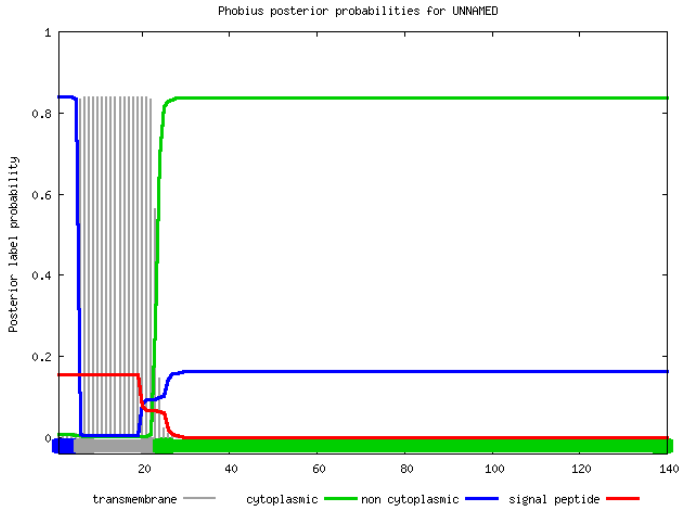


Figure A 3 Phobius plot for SAS1678.

Analysis of SAS1678 with phobius software showing high probability of protein not being secreted and contains transmembrane domain and is cytoplasmic in nature.

From gene DB 363 genes were described as hypothetical proteins from MSSA476. All 363 genes were screened for signal peptide, signal cleavage site, and non - cytoplasmic domains using different software's, Signal P 2 version (Fig A4), Signal P 3 version (Fig A5), Phobius (Fig A6), Spoctopus (Fig A7) and TMHMM.(Fig A8) Each of this software's is capable of identifying transmembrane domains, signal peptides and cytoplasmic or non- cytoplasmic regions.

Each of the 363 genes protein sequence was identified on Kyoto Encyclopedia of Genes and Genomes website (KEGG). Each of the protein sequence were individually tested and results were gathered and analyzed for all 5 software together to determine if protein is secreted or non –secreted with probability of more than 0.5 (probability of 1 being 100%). Each of the genes was also checked for signal cleavage site with the probability of more than 0.5 (50%) using Signal P 2version, Signal P 3 version, and Phobius software's.

SignalP-HMM result:

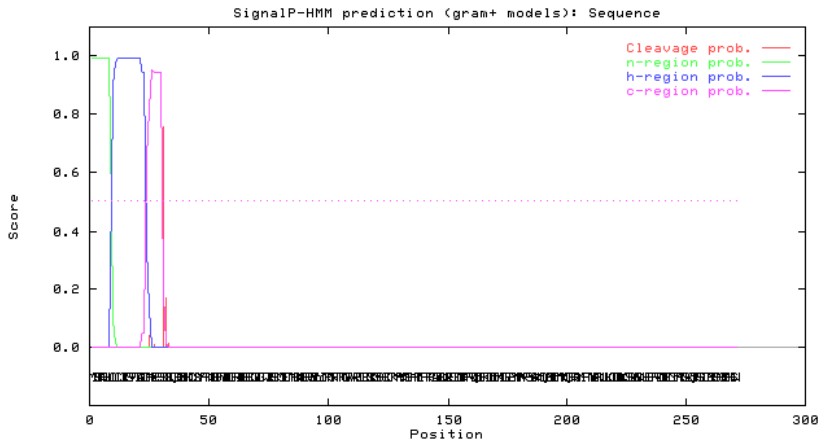


Figure A 4 Example of Signal P 2 version showing exported protein with signal cleavage site.

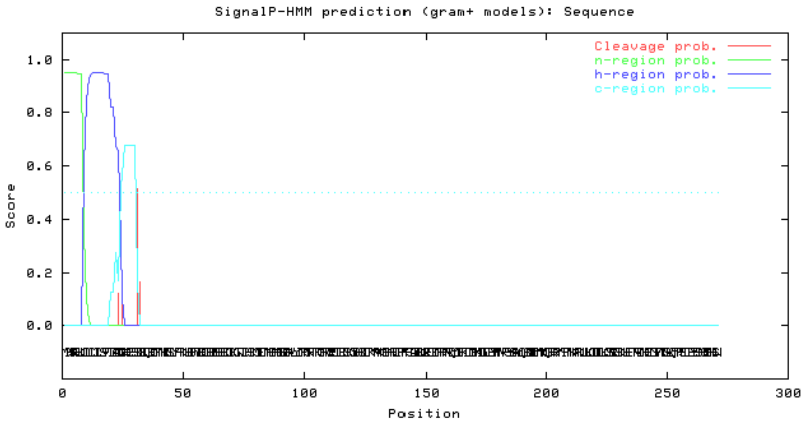


Figure A 5 Example of Signal P 3 version showing exported protein with signal cleavage site.

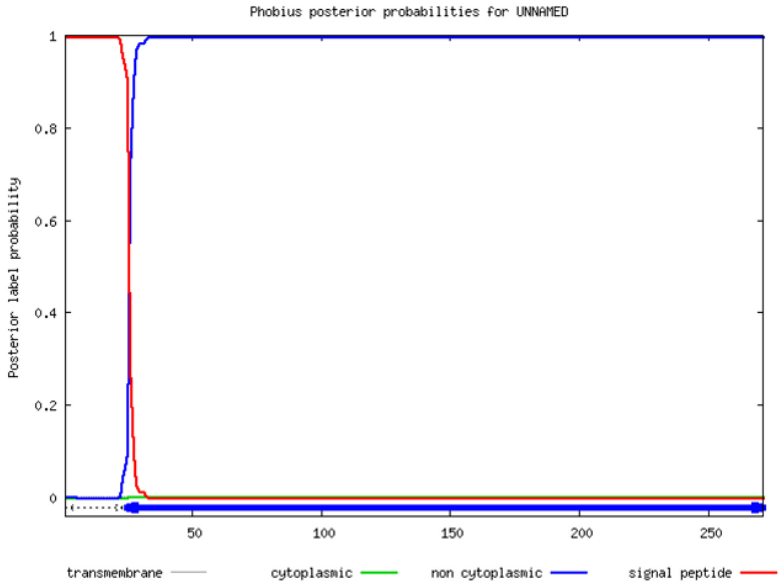


Figure A 6 Example of Phobius showing exported protein with signal peptide probability in red and non-cytoplasmic region in blue.

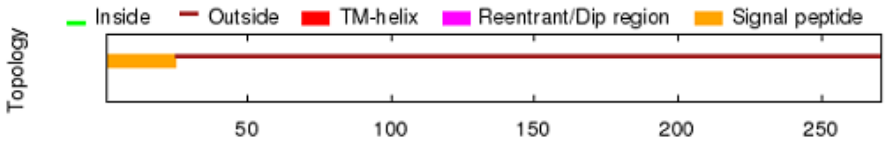


Figure A 7 Example of SPOCTOPUS software showing protein with signal peptide and non-cytoplasmic region (outside).

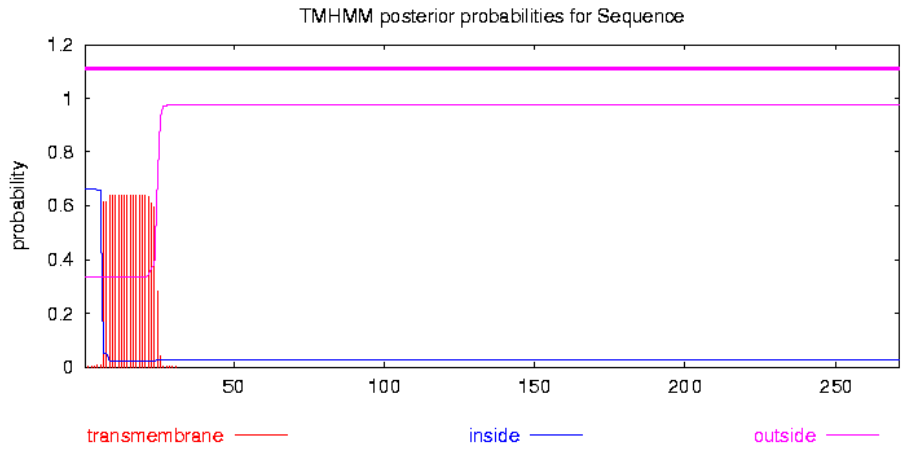


Figure A 8 Example of TMHMM software showing exported protein with transmembrane domain probability of more than 0.6 (60%).

Out of all 5 software we found TMHMM showing variable results compared to SignalP and Phobius software. TMHMM showed transmembrane domain probability which were not identified by Phobius or Signal P software. TMHMM software can identify transmembrane regions 98 – 99% accuracy but this software is not accurate in identifying transmembrane regions for proteins with signal peptide [77]. This may be the reason why results from Phobius and TMHMM varied so much.

Out of 363 genes, 47 genes were identified as hypothetical secreted proteins. All of the 47 gene nucleotide sequence was analyzed using BLASTn (NCBI) software to look for homologs or percent identity of these genes towards CA MRSA or HA MRSA.

To look for homologs or percent identity of these genes with CA MRSA or HA MRSA or both we located whole genome sequence of all *Staphylococcus aureus* strain on KEGG website, following the information and research articles associated with the search we classified *S. aureus* strains into CA MRSA, HA MRSA, or VISA strains. From this we identified homologs and

percent identity towards each of the 47 genes using BLASTn using nucleotide sequence and BLASTp with protein sequence of each gene.

Table A 1 Percent identity of test gene towards HA-MRSA.

All 17 genes showing 100% identity towards CA-MRSA are listed in this table also showing percent identity of these genes towards HA-MRSA.

Gene Number	**% Identity**
SAS0284	99% HA-MRSA
SAS0397	90% HA- MRSA
SAS0398	99% HA-MRSA, Newman
SAS0630	99% HA-MRSA
SAS0661	99% HA-MRSA
SAS0720	99% HA-MRSA
SAS0754	99% HA-MRSA, Lab
SAS0755	97% HA- MRSA
SAS0964	99% HA-MRSA
SAS1055	99% HA-MRSA, Newman
SAS1088	99% HA-MRSA
SAS2368	96% VISA
SAS2373	88% HA-MRSA
SAS2374	74% HA- MRSA
SAS2442	98% HA-MRSA
SAS2525	96% Col, Lab
SAS2189	99% HA - MRSA,

Out of 47 genes we found 24 genes that were secreted outside the cell, did not have any transmembrane domains, contains cleavage site, non- cytoplasmic, and unique to CA-MRSA with 100 % identity towards only CA-MRSA strains. Out of 24 genes 7 genes were already cloned in our lab, remaining 17 genes were again analyzed using protein BLAST to check for any homologous protein function that may give hint towards characterization of proteins under investigation.

Table A 2 BLASTp results with each of the test protein.

The BLASTp results showed some of the test genes similarity with proteins with known function or possible functions. Listed in the table are gene ID and number with possible function achieved by Blast P similarity search.

ID	Gene Number	Percent identity
1	SAS0284	Acid phosphatase activity, lipoprotein
2	SAS0397	Hypothetical exported protein
3	SAS0398	Hypothetical exported protein
4	SAS0630	N-acetylmuramoyl-L-alanine amidase, NLPC_P60
5	SAS0661	secretory antigen SsaA-like protein, LysM domain unique to MSSA476, Hypothetical exported protein
6	SAS0720	LysM family autolysin, CHAP domain-containing NLPC_P60, secretory antigen SsaA
7	SAS0754	Extracellular ECM and plasma binding protein
8	SAS0755	Extracellular matrix -binding protein
9	SAS0964	Lactococcin, low quality protein
10	SAS1055	Glycerophosphodiester phosphodiesterase
11	SAS2368	Unique to MSSA476, hypothetical protein
12	SAS2373	Uncharacterized lipoprotein
13	SAS2374	Tandem lipoprotein
14	SAS2442	Hypothetical protein

All 14 genes were characterized as hypothetical proteins; in addition to that few strain of *S. aureus* showing identity towards a lipoprotein or other possible function. But all genes showed 100 % identity towards hypothetical proteins alone with no possible function. Some proteins did show domains but looking at the individual domain we found a short sequence of the domain that showed similarity towards test protein.

Hypothetical Exported Proteins SAS2373, SAS2374 & SAS2375

Increased expression of SAS2373 & SAS2374 mRNA in the presence of keratinocytes cell lines shows factors associated with keratinocytes that may have triggered mRNA expression.

Both genes are located adjacent to each other and a lipoprotein SAS2375. clustalW alignment of all these genes shows similarity between all 3 genes. These proteins are hypothesized as tandem lipoproteins. Each lipoprotein consist of lipobox which aids in cleavage of prepostlipoprotein signal to release lipoprotein. A lipoprotein consist of positively charged aminoacids in the initial residues, presence of hydrophobic aminoacids between the positively charged residues and the lipobox, and presence of a Lipobox within the first 40 residues from the N-termninus with the consensus as [LVI][ASTVI][ASG][C]. Both genes SAS2373 and SAS2374 are negative for lipobox with little variation from that of gene SAS2375 (Fig A9).

```
sas:SAS2373      -MIHSKRLKLCLCLIILSVFIGACGMKKEESSKDKQIKENFNKILSLYPTKNLEDFYDKE
sas:SAS2374      --MHSKKLTLGICLVLLIILIVGYVIMTKANGQNAQIKDTFNQTLKLYPTKNLDDFYDKE
sas:SAS2375      MMIHSKRLRLWLYLVLLAVFISACGMKE-----DKQIKENFNKTLSLYPTKNLDDFYDKE
                   :***:* * : *::* ::* .  :       : ***:.**: *.*******:******

sas:SAS2373      GFRDEEFEKGDKGTWIIHSKMIIETNNSNMESRGMVLYINRNTRTTKGNFVVREITEDSK
sas:SAS2374      GFRDQEFKKGDKGTWIVNSGMNIQLKGGALKSREMVLYINRNTRTTKGYFIVGEITKDKK
sas:SAS2375      GFRDQEFEKGDKGTWIVDSEMVVELKDKKMESRSMVLYINRNTRTTKGNFIVRELWEDSK
                 ****:**:********:.* * :: :   ::** ************** *:* *: :*.*

sas:SAS2373      GY---------SHSKDTKYPVKMEHNRIIPTKPIADDKLRKEIENFKFFVQYGDFKDIND
sas:SAS2374      GYTHDKDKKGYTHDKDKKYPVKMEHNKIIPTKPIKDEKLKKEIENFKFFVQYGNFKDFKD
sas:SAS2375      GY---------AQSKDTKYPVKMEHNRIIPTKPIADDKLRKEIENFKFFVQYGDFKDIND
                 **         ::.**.*********:******* *:**:**************:***::*

sas:SAS2373      YKDGDISYNPNVPSYSAKYQLSNDDYNVKQLRKRYNIPTNKAPKLLLKGDGDLKGSSVGS
sas:SAS2374      YKNGDISYNPNVPSYSAKYQLSNDDYNIQQLRKRYDIPTKKAPELLLKGDGDLKGSSIGS
sas:SAS2375      YKDGDISYNPNVPSYSAEYQLSNNDYNVKQLRKRYDIPTKKAPKLLIKGDGDLKGSSIGH
                 **:**************:*****:***::******:***:***:**:**********:*

sas:SAS2373      KNLEFTFVENKEENIYFTDSVQYTPSEDTSYESNGISNKSW
sas:SAS2374      KDLEFTFVQNKRENIYFTDSVEFTPSEDTSYESN-------
sas:SAS2375      KNLEFTFVENKEENIYFTDSINFKPTK--------------
                 *:******:**.*******
```

Figure A 9 ClustalW alignment of genes SAS2373, SAS2374 & SAS2375.

The clustalW alignments shows high similarity between all 3 genes. Gene SAS2375 is positive for lipobox and is a lipoprotein. Whereas gene SAS2373 and SAS2374 consist of positively charged initial aminoacids residues and hydrophobic regions. At the site of lipobox, one aminoacids variation was observed with both genes SAS2373 and SAS2374.