MOLECULAR TYPING AND GENOME EXPRESSION PROFILING OF ORGAN- AND HOST-SPECIALIZED LINEAGES OF *STAPHYLOCOCCUS AUREUS* FROM BOVINE MASTITiS AND HUMAN INFECTIONS

BY

Kamaleldin B. Said
Department of animal Science
McGill University, Montreal
October, 2009

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

Staphylococcus aureus has become a major concern in public health and dairy industry due to the rapid evolution of host/organ specialized lineages adapted to humans and major food animals. However, the mechanism(s) of host- and organ-specialization as well as the distribution of dominant clonal lines of *S. aureus* is presently ill-defined. We hypothesized that coding repeat markers in the intragenic region of clumping factor A (*clfA*) gene would be capable of detecting and grouping adaptive clones and revealing predominant types and that whole genome expression of these clones in mammary mimicked reduced oxygen condition would potentially reveal subsets of gene(s) responsible for pathways and mechanisms of host-specialization and selection in the mammary gland. Thus, the suitability of *clfA* was explored. Results indicated that 80% of *clfA* analyzed had repeat-copies between 44 and 57. Furthermore, human isolates were polymorphic, while mastitis isolates were clonal. The repeats were stable during passages in milk, nutrient broth, and invasion of mammary cells showing suitability for typing. In addition, dominance of a clonal motif in mastitis implied organ-specific selection in the mammary gland. This was further examined in isolates from different organs in human patients and from bovine mastitis. The results showed significant correlation between the organs sampled and the length of *clfA*. Twenty out of the 23 sputum isolates had lower copy-numbers of 43-48, while 21 out of the 24 skin isolates had 55-63 copies. Moreover, sequence alignments and phylogenetic analysis placed isolates from different hosts and organs into respective clusters.

Subsequently, *clfA* was used as a typing marker for mastitis in a comparative experiment with pulsed-field gel-electrophoresis (PFGE) and *spa* typing, and identified genotype distributions and the dominant types across Canada. Interestingly, both *clfA* and PFGE had identical index of discrimination of 0.9. Two dominant PFGE lineage Groups A with 48.3% and D with 43.7 % represented Eastern and Western Canada strain types, respectively. These were further subdivided by *clfA* into four dominant subtypes X, Q, C, and Z which formed 82% and 43% of PFGE groups A and D, respectively. Thus, concordant with PFGE, *clfA* typing proved useful in rapidly profiling subpopulations with comparable discriminatory power.

Consequently, two isolates representing major human- and bovine-specific clones were used for whole genome expression after their invasion into mammary epithelial cells under normal and reduced oxygen conditions, using high-throughput genome qRT-PCR. In the mastitis isolate under normal oxygen, expression of *MerR, sigB*, VraS, YycG/YycF, araC, and *tetR* indicated adaptation, environmental sensing, binding, and protection. Coupling of fermentative metabolism to virulence was evident by upgulation of the catabolite control protein A (*ccpA)*, pentose monophosphate pathway, and down regulation of TCA cycle. Invasive phenotypes, potentially through *sarU* activation of *agr* and *in vivo* viability factors as staphopains and GntR operon, also occurred. However, under reduced oxygen it showed upregulation of fibrinogen-binding, *isd* operon, (*isdA, isdC, isdH)*, and *sdrH,* indicating aggressive binding phenotype. Irrespective of oxygen, strain Wright was less aggressive than the mastitis isolate and showed intense intracellular signalling potentially through *MerR* and *sigB*, limited toxins, and many hypothetical genes. Interestingly, under reduced oxygen, universal down regulation in metabolic and regulatory pathways occurred showing persistence of Wright strain. Taken together, the quick response of the mastitis isolate strongly suggested adapted phenotype to mammary cells and that reduced oxygen induced persistence properties. Future annotation of hypothetical proteins would potentially reveal a more precise network of regulatory subsets that are expressed in the mammary gland.

RESUME

Staphylococcus aureus est un problème majeur tant pour la santé publique que pour l'industrie laitière à cause de la rapide évolution de lignées spécialisées hôte/organe touchant l'Homme et les animaux à destiné alimentaire. Cependant, les mécanismes de spécialisation hôte/organe ainsi que la distribution de ses lignées clonales dominantes restent mal décrites. Nous avons postulé l'hypothèse que des marqueurs codants répétés dans la région intra-génique du facteur d'agglutination *clfA* permettraient de détecter et d'attribuer un groupe à des clones adaptatifs et d'en révéler ainsi les types prédominants et que l'étude de l'expression génomique totale de ces clones dans des conditions reproduisant celles de glandes mammaires privées d'oxygène pourrait potentiellement révéler des sous-groupes de gènes impliqués dans les mécanismes de spécialisation hôte et de sélection dans les glandes mammaires. Nous avons donc étudié le gène *clfA* et nos résultats indiquent que 80% des gènes *clfA* analysés avaient entre 44 et 57 copies-répétées. De plus, les isolats humains étaient polymorphiques alors que ceux provenant de mastite étaient clonaux. Les répétitions étaient stables durant les passages dans le lait, le bouillon nutritionnel et lors de l'invasion des cellules mammaires. Le gène c*lfA* s'est aussi révélé utile pour attribuer un groupe aux lignées hôte spécifiques et établir que la dominance d'un motif clonal dans une mastite implique une sélection organe-spécifique dans la glande mammaire. Ceci fut examiné plus avant dans les isolats de différents organes de patients humains et de mastites bovines. Nos résultats montrent une corrélation significative entre les organes étudiés et la longueur du gène *clfA*. Vingt sur les 23 isolats d'expectoration avaient un faible taux de copies (43-48) alors que 21 des 24 isolats de peau avaient entre 55 et 63 copies. Il est aussi important de noter que les alignements de séquence et les analyses phylogénétiques ont permis de distribuer les isolats provenant de différents hôtes et organes en des groupes respectifs.

Nous avons ensuite standardisé *clfA* comme marqueur pour la mastite au cours d'une expérience comparative utilisant des méthodes références, l'électrophorèse en gel pulsé et le *spa*, et ainsi identifié sa distribution génotypique ainsi que ses types dominants au Canada. Nos résultats montrent que *clfA* présente le même indice de discrimination de 0.9 que la PFGE. Deux groupes de lignage PFGE dominants (A avec 48.3% et D avec 43.7 %) représentent les types des souches présentes dans le Canada occidental et oriental

respectivement. Ceux-ci furent de plus subdivisés à l'aide de *clfA* en quatre sous-types. En conséquence, le typage par *clfA*, en plus d'être concordant avec ceux par PFGE et *spa*, s'est avéré utile pour établir le profil de sous-populations avec le même pouvoir discriminatoire.

Par la suite, deux isolats bien caractérisés représentant des clones majeurs spécifiques pour l'Homme et les bovins ont été utilisés afin de révéler l'expression génomique totale durant l'invasion des cellules épithéliales mammaires sous des conditions normales ou réduites en oxygène. Dans l'isolat de mastite sous condition normale en oxygène, l'expression des régulateurs *MerR, sigB*, VraS, YycG/YycF, araC, et *tetR* indiqua des fonctions d'adaptation, de fixation et de protection. Le couplage du métabolisme fermentaire avec la virulence fut mise en évidence par la présence de la protéine catabolite A (*ccpA)*, de la voie des pentoses monophosphate, et la régulation négative du cycle de Krebs. Considéré dans leur ensemble, les réponses rapides des isolats de mastite suggèrent fortement un phénotype adapté aux cellules mammaires. La description future du grand nombre de protéines hypothétiques pourrait potentiellement révéler un réseau précis de facteurs régulateurs qui sont spécifiquement exprimés dans la glande mammaire.

DEDICATION

I dedicate the best of this work to my mother and to the memories of my beloved father who taught us all that hard work isn't an option; rather, it is a necessity when you live for a strong purpose

I dedicate this work to my wife and the kids for understanding and allowing me to borrow from their precious time

I dedicate this work to those who dedicated their life to the progress of scientific thoughts for the Peace, Health, and Goodness of Mankind

ACKNOWLEDGEMENT

My sincere gratitude and appreciations goes to my committee members who guided me throughtout this work; my thesis supervisor Dr. Xin Zhao for his support, encouragement, and valuable advices, throughout the program, Dr. Arif Mustafa for his valuable advices, and discussions, and Dr. Donald Niven for answering my questions, and for the valuable discussions and directions.

I'm indebted to my family for their love, support, and patience without which this work would have never seen the light.

My gratitude goes to all our laboraotory members, to professors, friends and laboratory members of reproduction (specially Dr. Romain Rambrot for the excellent French translation), nutrition, breeding, and molecular biology in the Department of Animal Science for their collaborations, and to all other friends, students, faculty and members of the Macdonald Campus community for the good times, friendly interactions, and the vital discussions, and to those who contributed in anyway to the completion of this work.

CONTRIBUTION TO KNOWLEDGE

CHAPTER II: The study documented the development and evaluation of the suitability of the *clfA* locus as a rapid, highly specific, sensitive, and reproducible typing marker as well as its applicability for host-specific typing and differentiation of human and mastitis lineages of *Staphylococcus aureus.*

CHAPTER III: The study provided new perspectives on the organ-specificity of the *clfA* marker as applied to isolates of different human sites and bovine mammary gland. Acquisition of certain *clfA* lengths at specific host organs/sites has been a significant molecular correlate for host- and tissue-specializations in *S. aureus.* In addition, for the first time, concordance between repeat-based typing whole nucleotide sequence information confirms suitability of *clfA* use even without expensive sequencing.

CHAPTER IV: This study standardized the *clfA* typing through comparative analysis with gold standard methods pulsed-field gel electrophoresis and *spa* typing. The most significant advancement to science was the precise identification of dominant types as well as the further finer level subtyping within major PFGE types across Canada. Thus, *clfA* revealed, in an unprecedented study, the dominant repeat types as well as their evolution into subtypes at different geographic regions with the discriminatory power of that of the gold standard methods; however, *clfA* was meritorious in revealing the potential mechanisms of specialization and selection and in the rapidity, simplicity, and reproducibility of the system.

CHAPTER V: This comprehensive study has made the unprecedented contribution in revealing the subsets of genes, regulatory pathways, and the potential mechanisms underlying invasion of mastitis-associated and human-associated *S. aureus* into mammary cells. For the first time, preference of fermentative metabolism, over aerobic respiration, and its coupling to intense environmental sensing and virulence expression was revealed under normal growth with 5%CO_2. Specifically, potential activation of the *agr* through *sarU*, and the coupling of catabolite control protein pathways to virulence expression was revealed. In addition, effect of reduced oxygen on the cessation of metabolism and development of persistence properties were the major advancements.

STATEMENT ABOUT THE THESIS FORMAT

(Thesis office document)

This thesis has been written in the format of manuscripts submitted to scientific journals in accordance with the “Guidelines for Thesis Preparation, C: Manuscript-based Thesis” which states:

Candidates have the option of including, as part of the thesis, the text of one or more papers submitted, or to be submitted, for publication, or the clearly-duplicated text (not the reprints) of one or more published papers. These texts must conform to the "Guidelines for Thesis Preparation" with respect to font size, line spacing and margin sizes and must be bound together as an integral part of the thesis. (Reprints of published papers can be included in the appendices at the end of the thesis).

2. The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next. In order to ensure that the thesis has continuity, connecting texts that provide logical bridges preceding and following each manuscript are mandatory.

3. The thesis must conform to all other requirements of the "Guidelines for Thesis Preparation" in addition to the manuscripts.

The thesis must include the following:

1. a table of contents;
2. a brief abstract in both English and French;
3. an introduction which clearly states the rationale and objectives of the research;
4. a comprehensive review of the literature (in addition to that covered in the introduction to each paper);
5. a final conclusion and summary;
6. a thorough bibliography;

4. As manuscripts for publication are frequently very concise documents, where appropriate, additional material must be provided (e.g., in appendices) in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported in the thesis.

5. In general, when co-authored papers are included in a thesis, the candidate must have made substantial contribution in all papers included in the thesis. In addition, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. This statement should appear in a single section entitled "Contributions of Authors" as a preface to the thesis. The supervisor must attest to the accuracy of this statement at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to clearly specify the responsibilities of all the authors of the co-authored papers.

6. When previously published copyright material is presented in a thesis, the candidate must include signed waivers from the publishers and submit these to Graduate and Postdoctoral Studies with the final deposition, if not submitted previously. The candidate must also include signed waivers from any co-authors of unpublished manuscripts.

CONTRIBUTIONS OF AUTHORS

Four co-authored manuscripts submitted for publication are included in this thesis.

In all manuscripts, **Kamaleldin B Said** developed the original concepts, planned and co-ordinated the research, designed and carried out experiments, interpreted and analyzed the results, and wrote the manuscripts.

Authores of Manuscript 1 (Chapter II): Said KB, Ramotar K, Zhu G, and Zhao X.

Said KB designed and carried out all experiments, analyzed results, and wrote the manuscript. Zhu G assisted in experiment. Ramotar K and Zhao X assisted in experimental design and manuscript review.

Authores of Manuscript 2 (Chapter III): Said KB, Zhu G, and Zhao X.

Said KB desgined and carried out experiments, analyzed results, and wrote the manuscript. Zhu G assisted in experimental design, and Zhao X assisted in experimental design and manuscript review.

Authores of Manuscript 3 (Chapter IV): Said KB, Ismail J, Campbell J, Mulvey MR, Bourgault A-M, Messier S, and Zhao X.

Said KB designed and carried out experiments and analyzed results. Zhao X and Mulvey MR assisted Said KB in the experimental design. The other authors assisted Said KB in experiment and data analysis. The manuscript was written by Said KB and reviewed by the other authors.

Authores of Manuscript 4 (Chapter V): Said KB, Jones MB, Saeed AI, Dracheva T, Peterson S, and Zhao X.

Said KB initiated the original concept, prepared candidate *S. aureus* strains, designed, and carried out experiments, analyzed, and interprested results. Zhao X assisted Said KB in the experimental desgin, manuscript editions, and supervision. Jones MB and Peterson S assisted with expression experiments and result. Dracheva T and Saeed AI, assisted in data analysis. Manuscript was written by Said KB and edited by Zhao X.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AFLP..... Amplified Fragment Length Polymorphism

AFLP..... Amplified Fragment Length Polymorphism

CC..... Clonal complexes

CA-MRSA.....Community Acquired MRSA

Cp……...Crossing point of fluorescence signal

DMEM.....Dulbecco's Modified Eagle Medium

DNA.....Deoxyribonucleic acid

FBS.....Fetal Bovine Serum

ET.....Electrophoretic types

HA-MRSA.....Hospital Acquired MRSA

ID.....Index of discrimination

IMI…… Intramammary infections

MIC.....Minimum inhibitory concentration

MLST .....Multilocus sequence typing

MRSA.....Methicillin-resistant *Staphylococcus aureus*

MSSA.....Methicillin sensitive S.aureus

MSCRAMM.....Microbial-Surface-Component Recognizing Adhesive Matrix Molecules

MOI……. Multiplicity of infection

PFGE.....Pulsed field gel electrophoresis

PVL.....Panton valentine leukocydin

RT.....Repeat-types

SCC..... Somatic Cell Count

SCV.....Small Colony Variants

ST.....Sequence types

TCRS.....Two components regulatory systems

TRF.....Tandem Repeat Finder

TSB.....Trypticase soy broth

TR.....Tandem repeat

VNTR.....Variable-number tandem-repeat

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. GENERAL INTRODUCTION AND LITERATURE REVIEW

GENERAL INTRODUCTION

Mastitis has been one of the major and widely known economic problems of the dairy industries in the world. It is also an animal welfare as well as a public health issue due to concerns about residual antibiotics, resistant bacteria, and resistance genes in the food chain (Salisbury et al., 2002; Lee, 2003). Although several preventive measure, such as optimization of milking practices, culling, post-milking teat disinfection, and antibiotic therapy, have been practiced, to minimize the incidence of bovine mastitis (Zecconi et al., 2003), this disease continues to be a significant obstacle. Although many bacterial species cause mastitis, *Staphylococcus aureus* (*S. aureus*) has been one of the most common isolates. The invasive and chronic natures of *S. aureus* diseases and its highly adaptable mechanisms of pathogenicity make this species a difficult pathogen. It responds poorly to antibiotic therapies that are often unsuccessful in preventing or eliminating chronic udder infections. Although a number of attempts have been made to develop a vaccine against *S. aureus*, (Michie, 2002; Brouillette et al., 2002; and Carter and Kerr, 2003), the full protection desired against current and recurrent infections is still awaited. Thus, despite enormous efforts, the basic mechanisms of *S. aureus* pathogenicity, host specialization, and epidemicity are still quite elusive.

The recent years have been witnessing one of the major shifts in *S. aureus* disease epidemiology and pathogenesis in humans and animals. Evolution of community acquired methicillin resistant *S. aureus* (CA-MRSA) lineages with higher epidemicity and virulence have been increasingly reported in humans (Todd Webber, 2005; Vandenesch et al., 2003) and in animals worldwide (Weese 2005; Weese et al., 2005). A notable example is the changing epidemiology of C-MRSA from compromised patients to young adults who were otherwise healthy (Hidron et al., 2009). Current *S. aureus* research is mainly focused on two stage strategies, the pathogen and the clinically ill person. The zoonotic as well as zooanthroponotic (where strains could be transmitted to animals from humans) aspects which are of major concerns to animal scientists and public health officials (Epstein and Prince, 2009; Childs and Gordon 2009; Stephen et al., 2004) has

been largely ignored. Thus, the genotype distribution of mastitis-specific *S. aureus* and the subsets of genes and pathways underlying basic mechanisms of specialization are largely unknown.

Despite of the fact that *S. aureus* posses same gene content and genetic background in humans and animals (van Leeuwen et al., 2005), quite insufficient effort is made to understand the pathogenesis of *S. aureus* and its impact on dairy cattle. For instance, subclinical and clinical *S. aureus* mastitis surveillance and typing data, which are key tools to monitoring intramammary infections, have been poorly documented compared to its human counterparts. There have been substantial efforts in the development of new typing and diagnostic technologies for human associated *S. aureus*, including the use of selective culture media and real-time PCR assays, for faster detection. Nevertheless, no single *S. aureus* typing method is yet providing fully reliable information and discriminatory power needed (Struelens et al., 2009). Development of highly discriminatory typing markers that couple differentiation based on functional diversities of strains is critical in the characterization and definition of strains as host-specific lineages before subsequent use in comparative genomics and expression profiling. Recent comparative genome alignment and expression studies demonstrated that subtle nucleotide changes and differential expression, rather than acquisition of extrachromosomal elements, have been suggested to account for superior pathogenicity and host-specific functional diversity (Li et al., 2009; Highlander et al., 2007; Sivaraman et al., 2009). Thus, the global objective of this study was therefore to develop a *clfA*-repeat-based objective strain typing method for mastitis-associated *S. aureus* and to identify common genotypes and their distribution with the final goal of determining genome expression profiles of these types for subsets of genes and pathways underlying the basic mechanisms of host- and organ-specialization and selection in the mammary gland. We have aimed to provide new perspectives in the use of host- and organ-specific typing tools for precise differentiation of specialized types and their subsequent high-throughput genome wide expression profiling that will reveal a precise network of regulatory processes involved in the mammary gland microenvironment

LITERATURE REVIEW

1. Canadian dairy sector and economics of mastitis

The dairy sector is one of the largest agricultural industries in Canada. This industry operates under a supply management system, and provides a wide range of high quality dairy products. According to the Canadian Dairy Commission report (2006) available at (http://www.cdc.ca/cdc/index_en.asp?caId=87), the dairy sector contributed 11.5 billion CAD value to the Canadian economy and generated a net 4.84 billion CAD from farm cash receipts alone. The majority of dairy farms (81%) are located in Ontario and Quebec, while the remaining farms are in the Western and Atlantic provinces. In 2005-2006 dairy-year, there were 1.06 million cows in Canada, housed on 15,552 dairy farms, and delivered 75.57 million hectolitres of milk (Canadian Dairy Commission 2006 http://www.cdc.ca/cdc/index_en.asp?caId=87). Unfortunately, over decades, mastitis remained the most devastating disease in the dairy industry worldwide (Seeger et al., 2003) causing serious economic losses as high as $300-400 million CAD and $2 billion USD in Canada and the US respectively (Canadian Bovine Mastitis Research Network, 2006; National Mastitis Council, 2005) http://w3.aces.uiuc.edu/AnSci/USDA/NE-112/Background.shtml. Similarly, in New Zealand, reduction in milk production and quality as well as antibiotic treatment costs as a consequence of mastitis has been estimated at $300 million NZ annually (Denis et al., 2009).

2. Bovine mastitis

Mastitis refers to ‘inflammation of the mammary gland’, that generally occurs as a consequence of an inflammatory immune responses against intramammary (IMI) bacterial infections. The acute phase of the infection is characterized by physical, chemical, and usually bacteriological changes in the milk as well as pathological changes in the udder; these changes may become systemic in severe cases. Depending on the symptoms shown by the cow, mastitis can be clinical or subclinical. During the former, the udder is visibly inflamed, clots and clumps of blood appear in the milk, and the cow looks sick. On the other hand, during subclinical mastitis, both the cow’s health and milk “appear” normal, while milk production is reduced and somatic cell counts (SCC)

increased. Mastitis can also be chronic with long-lasting or recurrent infections. The chronic type is one of the most difficult types since the infectious agent becomes intracellular in the mammary gland, and hence, cannot be easily identified or controlled by antibiotic therapy. This dormant phase of bacteria (e.g., *S. aureus*) may become the reservoir of, sometime life-lasting, recurrent infections.

2.1 Types of mastitis

Based on the source of the infectious agent, mastitis is defined as *environmental mastitis* or *contagious mastitis*. It is called environmental mastitis if the source of the infection is the cow's environment such as water, soil, bedding, manure, air or feed, while contagious mastitis refers to infection transmitted from cow to cow or human to cow.

2.1.1 Environmental mastitis

The most common pathogens that cause environmental mastitis are *E. coli, Streptococcus uberis*, *Pseudomonas aeruginosa, Enterococcus faecalis* and *E. faecium*. Infections caused by *E. coli* are usually characterized by a very acute and sudden onset of mammary gland inflammation usually occurring around parturition and during early lactation. If not attended, the disease leads to severe local and systemic inflammations, and sometimes to death. However, the prevalence rates of mastitis significantly differ in different countries due to the climate and the complex nature of the disease. In Canada, among 3,149 cases of clinical mastitis, *E. coli* was one of the most frequently isolated bacteria associated with highest bulk milk SCC (Olde Riekerink et al., 2008). While Klebsiella clinical mastitis was less frequent than that of *E. coli* in Europe, these were of equal importance in the US (Barkema et al., 1998; Roberson et al., 2004). In the US, coliforms have been frequently identified in the warmer regions (Erskine et al., 1988). In New Zealand, coliforms were less frequent as mastitis-causing pathogens, but *Strep. uberis* was the main concern in both clinical and subclinical mastitis (McDougall, 1998).

As a consequence of environmental mastitis, many high producing cows are affected in a short period of seasonal time leading to death cases in the most severe situations (Burvenich et al., 2003). Nevertheless, there is currently sufficient evidence to believe that the severity of *E. coli* mastitis is mainly determined by cow factors as a

function of host neutrophils (Burvenich et al., 2003). Depending on herd management practices and hygienic conditions such as types and quality of water, beddings, soil, and stalls used, coliforms usually tend to prevail in dairy herds. Finally, prophylactic measures such as immunization programs with the *E. coli* J5 vaccine have been proved effective for control and prevention of coliform mastitis (Dosogne et al., 2002; LeBlanc et al., 2006).

2.1.2 Contagious mastitis

Contagious mastitis is usually caused by gram positive bacteria such as *S. aureus*, coagulase-negative staphylococci, and *Strep. agalactiae*. *S. aureus* is one of the most well known gram-positive pathogens that cause subclinical and clinical mastitis in dairy cows, and is usually associated with elevated SCC. Coagulase-negative staphylococci are widely known as normal commensals of skin of humans and animals. *Strep. agalactiae* is a gram-positive bacterium, that most often causes subclinical mastitis and elevated SCC (Keefe et al., 1997). On the other hand, *Mycoplasma,* pleomorphic bacteria which lack cell wall, has been classified as contagious mastitis pathogens, but development of the disease follows the environmental pattern (Bramley, 1992).

Contagious infection usually starts by binding of bacterial cells (individually or in groups) to host cells, followed by colonization (or invasion) and multiplication leading to pathological changes in the host cells (disease). In *S. aureus* mastitis, bacteria are usually transmitted from cow to cow through cloths, milking machines, air, or human contacts. Upon colonizing the teat canal opening, the pathogens migrate up the canal to reach the milk producing secretary alveolar cells in the mammary gland. Intracellular in the mammary gland, *S. aureus* undergoes significant changes in gene expression that leads to comprehensive changes in its physiology, resulting in dormant variants (Proctor et al., 2006). Contagious pathogens such as *S. aureus* and *Strep. agalactiae* are known to progressively increase somatic cell counts (Bramley, 1992). However, it also depends on the region and management system; in the tie-stall system, where the contagious transmission is easy due to physical proximity, *S. aureus* have been usually the most frequent isolate in clinical mastitis cases (Ericsson et al., 2009). Further, in Norway, it was the most frequently isolated bacteria from clinical mastitis samples, followed by

Strep. dysgalactiae (Reksen et al., 2006). Irrespective of the management system, the mode of transmission, epidemicity, and virulence properties of mastitis pathogens have been monitored by epidemiological typing strategies, which will be discussed later.

2.2 Public and animal health concerns of mastitis

Mastitis is both an economic as well as an animal health issue. The effect of first occurrence of pathogen-specific clinical mastitis on milk yield was studied in a comprehensive study comprising 3071 dairy cows caused by *Streptococcus* spp., *S. aureus*, other *Staphylococcus* spp., *E. coli*, *Klebsiella* spp., and *Arcanobacterium pyogenes*. It was found that clinical mastitis has significant impact on cow's health and milk production and mastitic cows often never recover their potential yield status. In primiparous lactating cows, *S. aureus*, *E. coli*, and *Klebsiella* spp. caused the most losses. In older cows, the above three pathogens, *Streptococcus* spp., and *A. pyogenes* caused the most significant losses (Grohn et al., 2004). Mastitis is also one of the main mortality causes in dairy cows. According to the Dairy USDA (2007) on the factors leading to mortalities in dairy cow, mastitis was reported as second only to the lameness or injury category reasons (Figure 1.1). A recent extensive analysis revealed that clinical mastitis significantly increased the risk of culling in dairy cows and mortality incidences, leading to both serious productivity and economic losses **(**Bar et al., 2008).

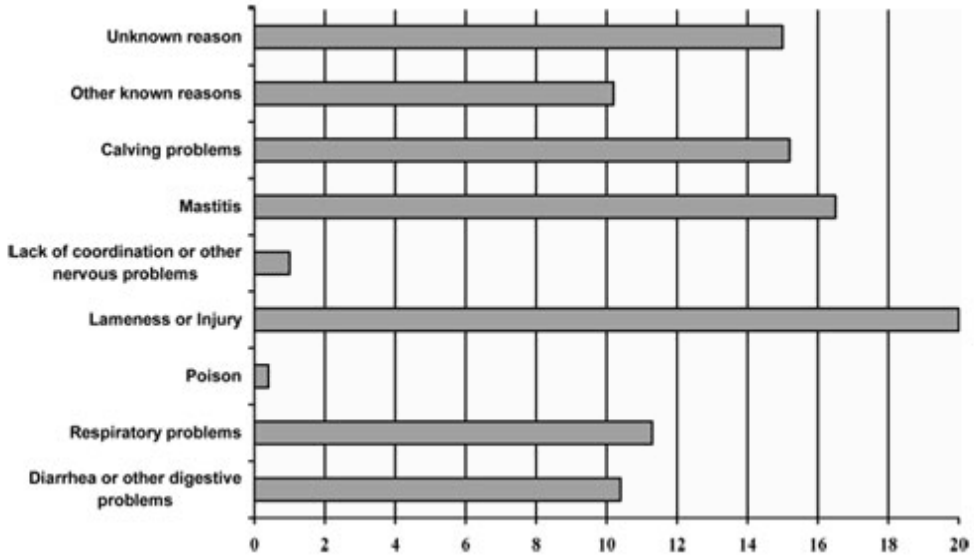


Figure 1.1. Percentage of cow deaths by producer-attributed cause (USDA survey) (Source: Dairy, 2007). Available at [http://www.thecattlesite.com/articles/1641/dairy-cow-mortality-a-growing-problem]

There have been different opinions within the medical, veterinary, and regulatory communities over whether the sub-therapeutic veterinary use of antibiotics would cause a significant risk for human health by compromising the antibiotic treatment of life-threatening conditions (Lammerding, 1998). The development of antibiotic resistance and foodborne illnesses are certainly considerable concerns in both health and economics costing about 6.9 billion US dollars annually (Allos et al., 2004). However, reports on high-frequency transfer of resistant determinants between human and animal strains are still rare with the exception of a few findings. For example, comparative genomic analyses have shown that the human MRSA252 strain uniquely shares multiple DNA sequence blocks with a global mastitis clones RF122 strain suggesting horizontal gene transfer between human and animal lineages (Brody et al., 2008). However, veterinary public health and preventive medicine play major roles in the fight against zoonotic diseases, those transmissible between humans and animals, by investigating the epidemiology of antimicrobial resistance and establishing risk assessment tools (Lather et al., 2001). Because of the usefulness of antimicrobial drugs in animals and humans, public health officials and veterinarians have increased efforts to reduce the indiscriminate use of antibiotics. In one report, they believe that one way to prevent antibiotic resistance is through outreach efforts that encourage and facilitate information flow from veterinarians to dairy producers who use specific antibiotics such as gentamicin without veterinary consultation (Cattaneo et al., 2009). In another report, the Board of Regents of the American College of Veterinary Internal Medicine has issued a consensus statement as a guide to veterinarians for the care and well-being of animals under their care with the protection of other animals and public health (Morley et al., 2005). The commonly used antibiotics in US are shown in Table 1.1 (Mathew et al., 2007).

Table 1.1 ANTIBIOTIC RESISTANCE IN FOOD ANIMAL BACTERIA (Mathew et al., 2007) 117

TABLE 1. ANTIBACTERIAL PRODUCTS APPROVED FOR USE IN LIVESTOCK IN THE UNITED STATES[a]

Drug	*Antibiotic family*	*Animals used in*	*May be used in feed*	*Used in human medicine*
Amoxicillin	β-lactam	B, D, P, S	No	Yes
Ampicillin	β-lactam	B, D, P, S	No	Yes
Apramycin	Aminoglycoside	S	Yes	No[b]
Arsanilic acid	Arsenical	P	Yes	No
Avilamycin	Orthosomycin	S	Yes	No
Bacitracin	Bacitracin	B, D, P, S	Yes	Yes
Bambermycin	Bambermycin	B, D, P, S	Yes	No
Carbadox	Quinoxaline	P, S	Yes	No
Ceftiofur	Cephalosporin	B, D, P, S	No	No[b]
Chlortetracycline	Tetracycline	B, D, P, S	Yes	No[b]
Cloxacillin	β-lactam	B, D	No	Yes
Colistin	Polypeptide	P	Yes	Yes
Danofloxicin	Fluoroquinolone	D	No	No[b]
Efrotomycin	Elfamycin	S	No	No
Enrofloxacin	Fluoroquinolone	B	No	No[b]
Erythromycin	Macrolide	B, D, P, S	No	Yes
Florfenicol	Phenicol	B, D, P	No	No[b]
Gentamicin	Aminoglycoside	B, D, P, S	No	Yes
Hygromycin	Aminoglycoside	P, S	Yes	No
Lincomycin	Lincosamine	B, D, P, S	Yes	Yes
Neomycin	Aminoglycoside	B, D, P, S	Yes	Yes
Novobiocin	Novobiocin	B, D, P	Yes	No
Oleandomycin	Macrolide	B, D	No	No
Oxytetracycline	Tetracycline	B, D, P, S	Yes	Yes
Penicillin	β-lactam	B, D, P, S	Yes	Yes
Pirlimycin	Lincosamine	B, D	No	No[b]
Polymyxin	Polypeptide	B, D	Yes	Yes
Roxarsone	Arsenical	B, D, P, S	Yes	No
Spectinomycin	Aminocyclitol	B, D, P, S	No	Yes
Sulfachlorpyridizine	Sulforamide	B, D, S	No	No[b]
Sulfadimethoxine	Sulforamide	B, D, P, S	No	No[b]
Sulfaethoxypyridazine	Sulforamide	B, D, P, S	No	No[b]
Sulfamethazine	Sulforamide	B, D, P, S	Yes	No[b]
Sulfathiazole	Sulforamide	B, D, S	Yes	No[b]
Tetracycline	Tetracycline	B, D, P, S	No	Yes
Tiamulin	Diterpene	S	Yes	No
Tilmicosin	Macrolide	P, S	Yes	No[b]
Tylosin	Macrolide	B, D, P, S	Yes	No[b]
Tulathromycin	Triamilide	B, D, S	No	No
Virginiamycin	Streptogrammin	P, S	Yes	No[b]

[a]Adapted from CVP, 2006; Guardabassi and Couravalin, 2006; FDA, 2006b.
[b]Closely related analogs are used in and are of importance to human medicine.
B, beef cattle; D, dairy cattle; P, poultry; S, swine.

3. *Staphylococcus aureus* in human disease

Staphylococcus aureus has been one of the most important and fastest re-emerging pathogens in human medicine. It is commonly found associated to human skin and anterior nares with carriage rates of about 20–60% in the general population (Kluytmans et al., 1997; Jernigan et al., 2003). It is well known to cause a very wide

range of infectious diseases ranging from common simple skin infections such as boils, pimples, impetigo, styes to more severe infections such as endocarditis, necrotizing pneumonia and necrotizing fasciitis (flesh eating disease). *S. aureus* is a major cause of skin and surgical-site infections and is one of the most common causes of healthcare-associated infections. As it is a major cause of mastitis in the dairy cows, it is also an important cause of puerperal mastitis and cesarean section infections in women in about 25 to 50% of cases (Sweet et al., 2001). More importantly, the rapid increase in methicillin resistance and evolution of hypervirulent strains in different lineages has been one of the most difficult tasks that faced global health care systems. For instance, in the year 2005, a single lineage of *S. aureus* caused more invasive diseases than the combined rates of diseases caused by bacterial species with most transformable genomes; namely, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Neisseria meningitidis*, and *Haemophilus influenza*. *S. aureus* mortality rates well exceeded that of the HIV (Bancroft, 2007, Klevens et al., 2007).

4. Epidemiology and genotyping

Several typing methods have been used in classifying *S. aureus* in human infections and in bovine mastitis, including phage typing, pulsed field gel electrophoresis (PFGE), binary typing, multilocus sequence typing (MLST), as well as phenotypic traits (Roberson et al., 1998; Enright et al., 2002; Rabello et al., 2007; Zadoks et al., 2002). Multilocus sequence typing grouped *S. aureus* into five main clonal complexes (CC), namely, CC8, CC30, CC5, CC22, and CC45 (Enright et al., 2002; Feil et al., 2003; 2004; Robinson and Enright, 2004). These complexes were represented in three major and two minor clusters of amplified fragment length polymorphism (AFLP) (Melles et al., 2004). Furthermore, high-throughput genotyping of human and animal isolates has consistently shown a common background and virulence gene content, but mastitis-specialized *S. aureus* belonged to distinct clusters (van Leuween et al., 2005). Pulsed field gel electrophoresis has been a gold standard for typing. It has an excellent discriminatory power, and considered adequate for closely related strains. More recent approaches include multilocus variable-number tandem-repeat (VNTR) typing methods. These have been suggested as useful additions to major molecular methods such as PFGE and DNA-

arrays. The first typical VNTR system for *S. aureus* was developed by Sabat et al., (2003) and used PCR amplicons of repeats from *sdr, clfA, clfB, ssp* and *spa* genes. However, single-locus repeat-based markers are potential approaches as alternatives to major methods. For instance, *spa*-typing has been one of the most successful single genetic markers for typing *S. aureus* (Koreen et al., 2004).

Application of epidemiological typing as a new tool for disease determinants, recognition of the multifactorial nature of mastitis and consideration of subclinical conditions, integration of the disciplines of veterinary medicine and animal science, understanding of nutritional influences (Goff, 2006), and the focus on science-based farm management have been some of the major advances in dairy health over the last two decades (LeBlanc et al., 2006). The different modes of transmission of the infectious agents within and between cows and human constitute the principle portals for the initiation of specialized clones. Epidemiological typing renders significant information available about local and global strain profiles, epidemicity, and virulence properties that can be utilized for control strategies. Although there has been a considerable progress in molecular typing of *S. aureus* from human infections, this has been slow for strains from major food animals even though several reports cited the significant similarity in genomic background and gene content of strains from different host species. Because *S. aureus* poses a risk of zoonotic infections, it is reasonable to suggest parallel surveillance programs in farm animals.

In recent years there has been a rise in *S. aureus* subtypes with elevated virulence and epidemicity in human and in dairy cows. In the dairy industry, the emergence of a hyper-virulent strain, namely ET3-1, a subtype of the common mastitis ET3 clone (Guinane et al., 2008) is a major concern. This specific subtype is highly prone to the acquisition of resistance determinants (Sung and Lindsay, 2007) posing a potential risk of new emergence of zoonotic infection (Juhasz- Kaszanyitzky et al., 2007; van Loo et al., 2007). Similarly, there has been a devastating clonal expansion and diversification of a subset of community-acquired methicillin-resistant *S. aureus* (CA-MRSA) isolates classified as USA300 (Kennedy et al., 2008; Klevens et al., 2007). It is well known that only a limited number of dominant clones are responsible for the majority of *S. aureus* infections, particularly in mastitis (Kapur et al., 1995; Fitzgerald et al., 1997, Smith et al.,

2005). Nevertheless, despite the clonality of the species genome based on both house-keeping and more polymorphic genes (Feil et al., 2003; Kuhn et al., 2006), how do the hypervirulent subsets of *S. aureus* evolve and what are the basic mechanisms underlying epidemicity and host- and tissue-specialization in different lineages has been quite elusive. Thus, there is an urgent need for objective strain-subtyping markers that couple epidemiological investigations with pathogenesis and colonization potentials. One of these markers has been the tandem intragenic repeats found in prokaryotic genomes. The universality, sequence characteristics, structural uniqueness, differential expression patterns, their direct involvement in virulence, and their demonstrated evidence in niche adaptations, are some of the properties that render repeats of the surface proteins as potential candidate markers that would reveal epidemic and colonization potentials.

5. Mechanisms of *S. aureus* invasion into host cells

Pathogenic bacteria exploit host cells for their own uptake into the cells by either of the two major types of induced mechanisms, the "zipper" or the "trigger", using host signal transduction pathways and cytoskeletal rearrangements (Finlay and Cossart 1997). The trigger mechanism is characterized by dramatic host membrane ruffling; whereas, the zipper mechanism is typified by phagocytic cup formation and requires fimbriae-mediated adhesion. Microbial pathogens using the zipper mechanism to enter nonprofessional phagocytes (e.g. epithelial cells) express fimbrial adhesins to bind host surface receptors involved in cell-to-cell adhesion, e.g., integrins. The expression of these adhesins leads to the formation of a vacuole that engulfs the bacterium through a 'zippering' process in which relatively modest cytoskeletal rearrangements and membrane extensions occur in response to engagement of the adhesin receptors (Cossart et al., 2004).

The invasion of *S aureus* into bovine mammary epithelial cells is a receptor-mediated endocytosis process and requires eukaryotic nucleic acid and protein synthesis and bacterial protein synthesis in the invasion process (Almeida et al., 1997). Thus, *S. aureus* invasion into mammary epithelial cells is dependent on the trigger mechanism. Several models have been described on the uptake machinery of *S. aureus* by bovine mammary epithelial cells, MAC-T. For example, Bayles et al., (1998) established that *S.*

aureus induce its internalization by a mechanism involving membrane pseudopod formation. Since each step in microbial infection is controlled by simultaneous expression of subsets of genes influence by global regulators, a model was suggested for the function of the *agr* locus in internalization, intracellular persistence, and dissemination (Wesson et al., 1998) as shown in Figure 1.2.

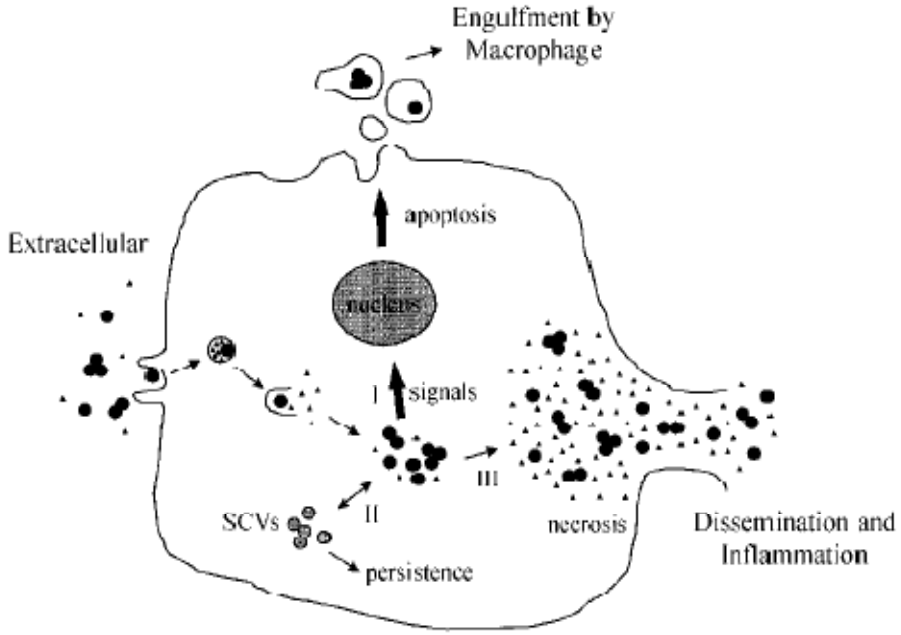


Figure 1.2. Model of Agr regulation during invasion. In an extracellular environment, levels of autoinducer peptide is low and *S. aureus* (solid circles) expresses cell surface-associated adherence factors that mediate uptake. Once internalized, a rapid accumulation of autoinducer within the endosome causes a shift to the expression of exoproteins that damage endosome and release the cocci. Once *S. aureus* resides in the cytoplasm of the host, three possible outcomes were proposed: (i) the induction of apoptosis, (ii) the formation of small colony variants (hatched circles), or (iii) lysis of the host cell. (Wesson et al., 1998).

Consequently, Shompole et al., (2003) confirmed that a biphasic intracellular induction in *agr* system within endosomes results in their escape into the cytoplasm following lysis of the endosomal membrane by *agr* induced toxins and products (Shompole et al., 2003). These observations indicate that at least two different subsets of gene classes that are under the controlled of two different regulatory loci, participate in the process leading to invasion. The initial extracellular attachment of bacteria to epithelial cell surface temporarily down regulates *agr* system resulting in significant reduction in metabolic pathways at the expense of increased expression of adhesion genes

such as *clfA* and FnBPs from alternative regulatory genes *sigB* and *sar* (Proctor et al., 2006; Vaudaux et al., 2002, Wolz et al., 2002). Following Wesson et al. proposal, the exact molecular interactions between the coccus and host cell was elucidate by Sinha et al., (1999) where adhesins initiated specific interaction through a fibronectin (or fibrinogen/fibrin) bridge between the coccus and host α5ß1 integrins leading to uptake (Sinha et al., 1999) as shown in Figure 1.3.

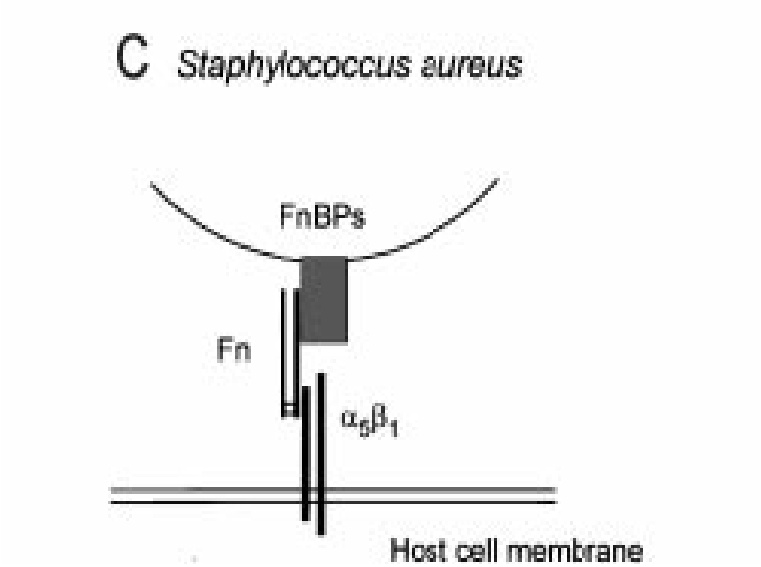


Figure 1.3. Interaction of *S. aureus* with host cell α5ß1 through a fibronectin bridge (Sinha et al., 1999).

6. Potential mechanisms for *S. aureus* host-specialization

1.1 Coding tandem-repeats

Based on numerous findings, intragenic repeat-mediated adaptation is one of the potential mechanisms underlying *S. aureus* mammary gland-specialization and evolution of mastitis-specific lineage. It is known that variations in phase of expression of a protein or in an antigenic repertoire of a bacterial strain generate corresponding diversity in phenotype of an otherwise clonal bacterial population. In one mechanism, protein expression is affected in which upstream repeat-triplets (usually in the promoter region) induce frame-shifts that act as ON/OFF switches to the phase-variable protein(s)

expression (van Belkum et al., 1997, 1998a, 1998b). In another mechanism, coding tandem repeats which do not alter the reading frame with copy number, are usually long and are in-frame for translation. These repeats code for the same proteins with different amino acid lengths thereby generating structural modifications that affect the function or antigenicity of the protein (Jordan et al., 2003).

S. aureus possess a large number of redundant cell surface adhesins in addition to virulent associated toxins, capsule, and exoproteins that are regulated through networks of complex regulatory genes (Figures 1.4 and 1.5). These are encoded by the species-specific variable genomic domain (*ori*C *environ*) which is responsible for staphylococcal-speciation (Takeuchi et al., 2005; Lindsay et al., 2006; Dordet-Frisoni et al., 2007), and contains genes for adherence and invasion mediated by specialized structures called microbial-surface-component recognizing adhesive matrix molecules" (MSCRAMM) (Foster and Hook, 1998) such as *spa* and *clfA*. The unique feature of these proteins is the presence of long tracks of hypervariable coding-tandem repeat in their C-terminal ends (Figure 1.4). In some MSCRAMMS such as fibronectin binding proteins (FnbPs), repeats are directly involved in ligand-binding; however, in other MSCRAMMs such as clumping factors A and B (encoded by *clfA* and *clfB*), the function of the repeat domains is still not clearly defined (Foster, 2005). There is accumulating evidence to believe that the repeat diversity in *clfA* and other similar proteins is a potential strategy, in a clonal background, for fitness in the host-adapted lineages of *S. aureus*.

The repeat sequence and structural organization within the *clfA* gene shows several clues for its role in adaptation. The repeat region (R-domain) is present at the downstream of the niche-specific ligand-binding domain A. It is composed of Ser-Asp dipeptide repeats [(18bp degenerate repeats that follow the consensus, [GAP (P is a pyrimidine), TCN (N is any base), GAP, TCN, GAP, AGP]. The R domain connects the cell-wall-spanning domain W to the ligand-binding A-domain (Figures 1.4 and 1.6) (Hartford et al., 1997, Foster and Hook, 1998). Some *in vitro* studies have suggested the importance of the repeat region. For example, a total of 72 residues were required between domain A and the conserved LPxTG sorting signal (site used for anchoring the protein to the cell wall) (Schneewind et al., 1995) to allow optimum clumping in fibrinogen. A stepwise decrease in this distance from 72 to 4 had resulted in a gradual

reduction in bacterial cell-clumping titers. In addition, mutants with 40 residues or less showed decreased binding of anti-ClfA serum and fibrinogen coated plastic surfaces, while further lower levels of binding were seen with null mutants of R region altogether (Hartford et al., 1997). Further, the number of repeats in the *clfA* has been shown to affect the adherence and clumping titers of the cocci *in vitro* (Risley et al., 2007).

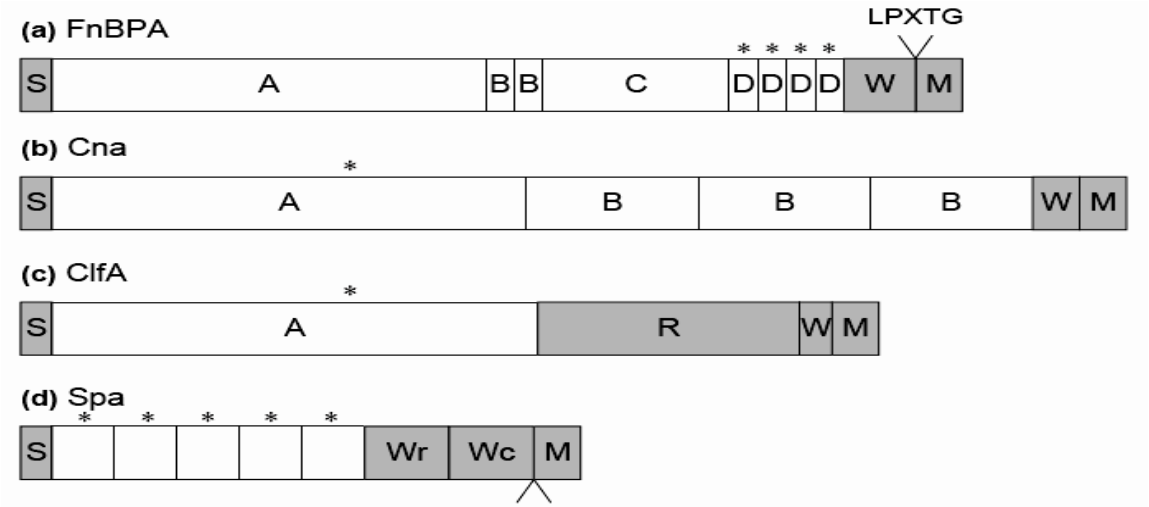


Figure 1.4. Structural organization of MSCRAMMs. S = signal peptide sequence, R = SD repeats, W = wall spanning region, M = membrane spanning region and positively charged residues. A, B, C, D domains and the LPXTG sorting signal domains are shown. Asterisks indicating ligand binding (Adapted from Foster and Hook 1998)

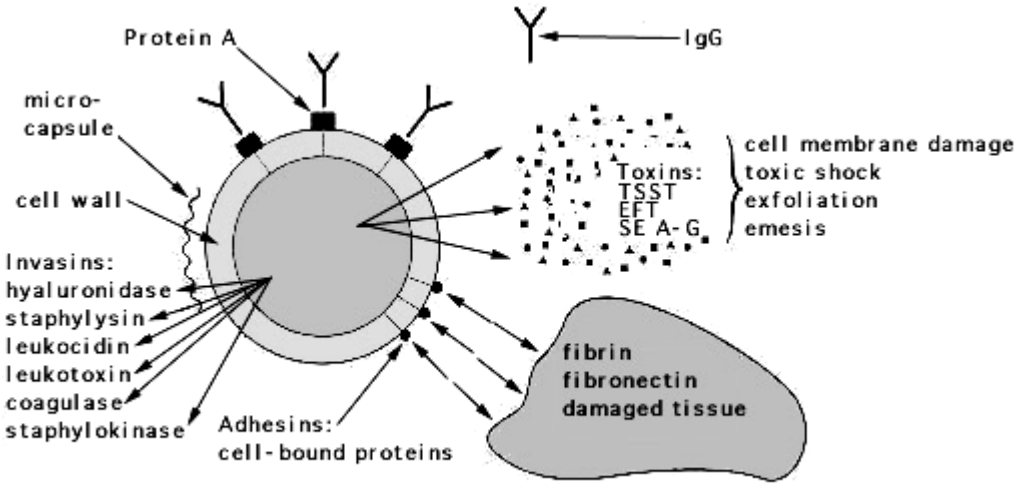


Figure 1. 5. Virulence determinants of *Staphylococcus aureus.* (Todar 2005). http://www.textbookofbacteriology.net

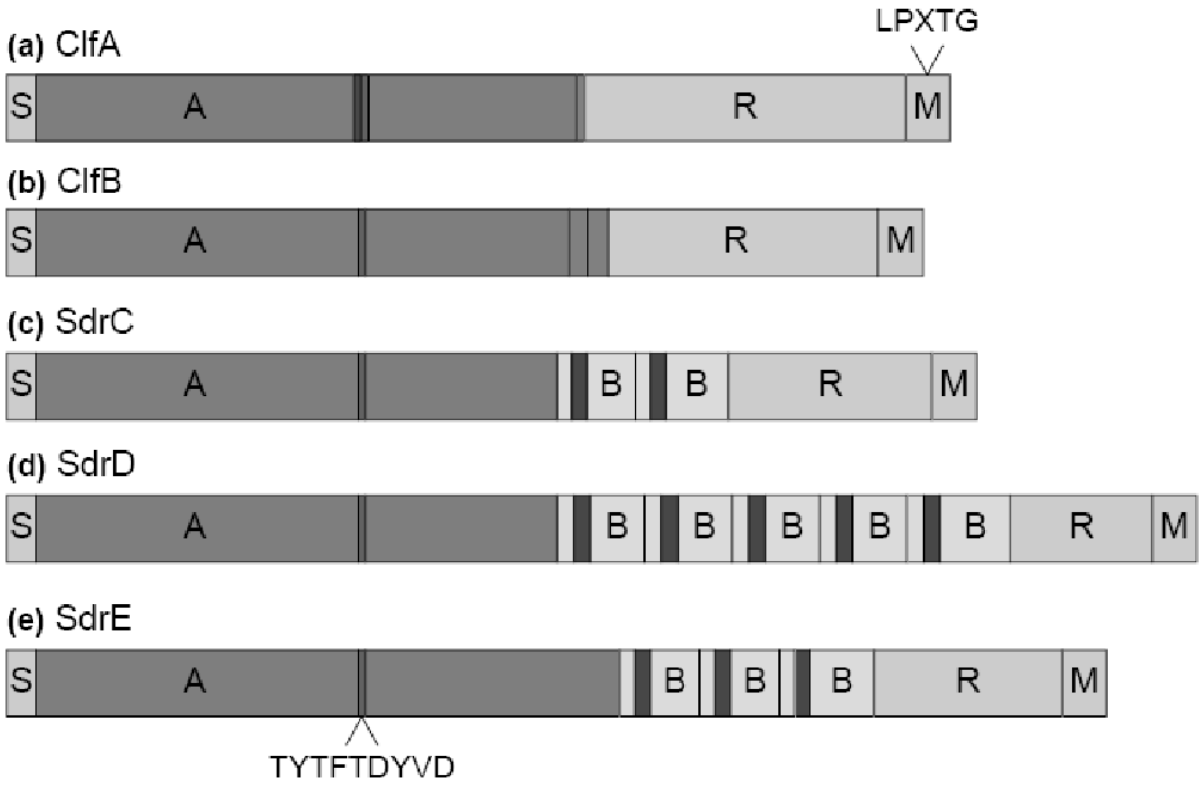


Figure 1.6. The ClfA–Sdr protein multigene family: (a) ClfA, (b) ClfB, (c) SdrC, (d) SdrD and (e) SdrE. S = the signal sequence; M = the LPXTG motif, and positively charged residues; R= Ser–Asp dipeptide repeats; A= the A-domain; and B = B repeats. The bold bars in repeat regions of Sdr C, D, and E = EF-hand loops within each repeat; in ClfA this loop is in A domain, and overlaps the TYTFTDYVD motif (thin bars in A regions) (Source: Foster and Hook, 1998).

Recent findings support the notion that clonal populations of bacteria respond to changes in the host microenvironment by minimal genetic change that can alter the virulence phenotypes. For instance, a significant reduction in experimental mouse mortality was attributed to only a very few point mutations in specific genes, including adhesion genes such as *clfA* (Kennedy et al., 2008). Moreover, several *in vivo* experiments indicated the importance of repeats in virulence. For instance, a reduction in the copy numbers of a coding TR within the α C-protein of *Streptococci* (from 9 to 1)

renders it unrecognizable by the anti-α C-protein antibodies thereby enhancing the virulence of the cocci (Madoff et al., 1996; Gravekamp et al., 1998). Similarly, the mycoplasma *vlp* gene undergoes size variations by intragenic expansion or contraction of a 3' region containing TRs; longer repeats mask antibody (Citti et al., 1997) and complement (Simmon et al., 2004) binding sites on the wall-less mycoplasma. In addition, the MSCRAMMs mediate host-pathogen interactions by niche-specific ligand-binding specificities (Projan and Novick, 1997, Foster and Hook, 1998). Finally, each staphylococcal species encodes a species- or strain-specific MSCRAMM with limited identity to other species (Ben Zakour et al., 2008), a property useful for their differentiation (Dordet-Frisoni et al., 2007).

6.2 Differential expression of adhesion genes

The mode of expression of ClfA in comparison to other similar proteins, such as ClfB, makes it an ideal tool for studying mammary gland-associated strains. The ClfB is expressed during the exponential growth phase under increased oxygenation; however, this protein is digested by the stationary phase proteases (Ni Eidhin et al., 1998; McAleese *et al*., 2001). This restriction in activity only at early phase and the sensitivity to proteases question its role in intracellular microenvironments such as IMI where variants such as the small colony types (SCV) are known to dominate (Proctor et al. 2006). On the other hand, a significant ClfA activity has been shown in this niche by IMI challenge of mice immunized with anti-ClfA antibody (Tuchscherr et al., 2008). Similarly, enhanced *clfA* and *spa* (another tandem repeat containing gene codes for protein A) expression was detected under bovine-mammary mimicked low oxygen condition, irrespective of strains used (Ster et al., 2005). Thus, ClfA is constitutive, independent of the Agr-system, abundant in deep-infections (Wolz et al., 1996, 2002; Josefsson et al., 1998), and mediates virulence even in absence of fibrinogen (Palmqvist et al., 2004a). In addition, ClfA, not fibronectin binding proteins, has been shown to account for the intracellular load of the cocci (Ahmed et al., 2001), and that ClfA was recently found as one of the genetic requirements for abscess formation and persistence (Palmqvist et al., 2004b; Cheng et al., 2009). Finally, ClfA has been found to have antiphagocytic properties comparable to that of Spa (protein A), and that could occur

independent of its ligand, fibrinogen (Higgins et al., 2006). The mechanism of action of ClfA is also similar to that of Spa where it cloaks the coccus surface preventing phagocytosis as shown in Figure 1.7 (Foster, 2005). However; SpA interaction in cows is not clear; it formed insoluble complexes with serum IgGs from Guinea pig and mouse only, but not from cow, goat, sheep, horse or chicken implying a different mechanism in hoofed animals (Atkins et al., 2008). Potentially, the spa's function is compensated by ClfA in these animals.

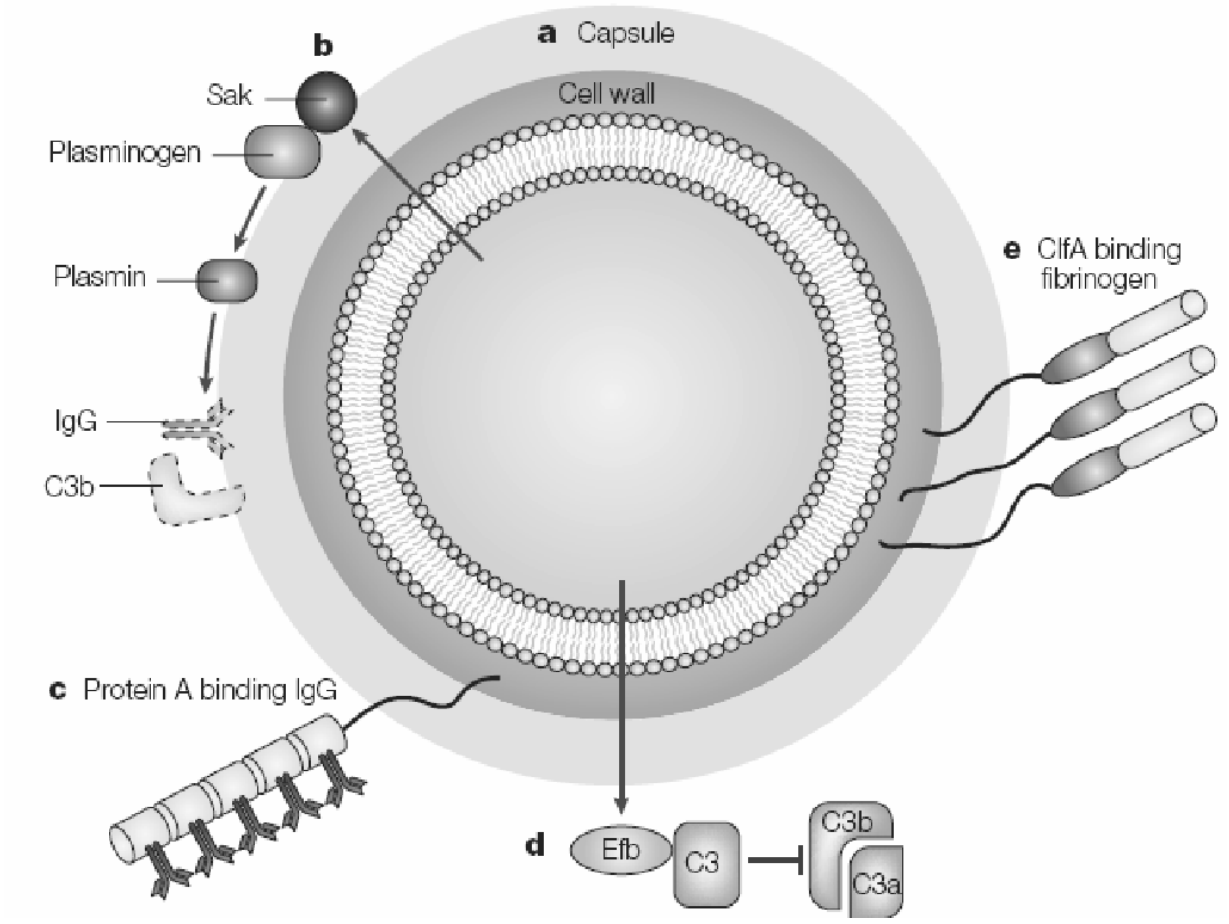


Figure 1.7. The mechanisms by which *S. aureus* avoids opsonophagocytosis. (Foster 2005). (**a**) Capsular polysaccharide prevents neutrophil access to bound complement and antibody. (**b**) Extracellular staphylokinase (Sak), activates plasminogen and cleaves IgG and C3b.(**c**) Protein A binds up to 5 immunoglobulin G (IgG) Fc-binding domains.(**d**) Fibrinogen-binding protein (Efb) binds complement factor C3 and blocks its deposition

on the bacterial cell surface. Complement activation beyond C3b attachment is prevented, thereby inhibiting opsonization. (**e**) ClfA, which binds the γ-chain of fibrinogen, and fibrinogen coated cocci are protected.

6.3 Differential expression of core-chromosomal genes

One potential pathogenic strategy of *S. aureus* is the adapted expression of core genome virulence determinants, rather than acquisition of mobile elements. Unfortunately, this is poorly documented in mastitis and only a few recent reports exist for human strains. Differential expression of *S. aureus* key virulence determinants, namely the phenol soluble modulins and alpha toxins, and not the acquisition of more virulent factors by horizontal transfer was responsible for the increased virulence in USA300 and in its progenitor USA500 (Li et al., 2009). Interestingly, expression patterns of these determinants in MSSA and MRSA isolates of the same type did not differ supporting the finding that methicillin-resistence, although it may aid in transmission, does not have an impact on virulence per se nor does it affect strain fitness (Bubeck Wardenburg et al., 2008; Li et al., 2009). Furthermore, CA-MRSA strains with identical genotypic and phenotypic properties as well as similar clinical characteristics have differed in their acquisition of the panton-valentine leukocidin gene (Zhang et al., 2008). Evidently, the strong association between the acquisition of genetic elements and the evolution and spread of hyper-virulent types has been changing in recent years (Rossney et al., 2007). These findings indicate that adaptive expression through unique quorum-sensing and microenvironmental conditions that stimulate unique regulatory pathways plays an important role in virulence.

7. Mammary oxygen impact on Staphylococcus aureus virulence regulation

It is known that *S. aureus* is highly adaptive to different sites in humans and animals and responds to environmental conditions and host intracellular milieu, which differentiates staphylococci (Josefsson et al., 2008), by modulating the global regulatory genes. These conditions control the two-component regulatory systems (TCRS) through a DNA-binding family of transcriptional regulators called SarA and their homologues, as well as the sigmaB factor. The transcriptional factors respond to host cell intracellular

signals such as oxygen content (Chan et al., 1998) leading to modulation of staphylococcal accessory gene regulator, *agr* system, as well as downstream target genes (e.g., *hla* and *spa*) (Cheung et al., 1997; 2008). It has been shown that insertion of a tampon raises vaginal oxygen from nearly anaerobic to aerobic conditions thereby activating *S. aureus* toxic shock syndrome in women (Wagner et al., 1984). Consequently, reduced oxygen has been shown to upregulate another two-component regulatory system knonw as staphylococcal respiratory response srrAB whch represses virulence factors such as TSST-1, and Spa (Yarwood et al., 2001). For these reasons, a better understanding of the regulation of the virulence would strongly require knowledge of the key conditions that *S. aureus* encounters inside the mammary gland that affect colonization and persistence capacities. Oxygen is one of these important conditions that would potentially alter the entire virulence process in this species. Oxygen is low in the uninflammed mammary gland reaching about 23 mmHg; however, that level has been found further much reduced in *S. aureus* infected glands accounting to only 1.3 mmHg (Mayers et al., 1988). Several regulatory systems have been studied in *S. aureus* (Novick et al., 2003; Bronner et al., 2004; Cheung et al., 2008) and many effects of environmental conditions have been shown to affect virulence (Cosgrove et al., 2007; Brown et al., 2003). However, most of the previous studies used laboratyory passed strains that are lilely to have differenes in expression (Blevins et al., 2002; Novick et al., 2003). In addition, since different cell types have different intracellular conditions, the response of *S. aureus* to mammary cell is expected to be unique during bovine mastitis situations. Thus, the low oxygen level is much likely to have a significant impact on the virulence processes and may induce alternative global regulatory routes in the mammary gland.

8. Quorum-sensing and global regulation

Quorum-sensing is a multi-cellular behavior that synchronizes the activity of bacterial population by releasing, detecting, and responding to small signal molecules called pheromones or autoinducers. The concentration of excreted autoinducers in the immediate microenvironment of the bacteria is directly proportional to their “density” and that regulates group behavior on a population-wide scale. This is accomplished when the chemical information in the signal molecules is integrated, processed, and transduced to

control gene expression through global regulatory genes at specific levels. At one level, expression is controlled by TCRS for which the staphylococcal accessory gene regulator, *agr,* is the main quorum sensing mechanism in *S. aureus*. It has two sets of operons controlled by two promoters P2 and P3. The latter controls levels of the *agr* effector molecule, the RNAIII that encodes membrane damaging factors and toxins such as δ-hemolysin (Figure 1.8) (Yarwood and Schlievert, 2003). The former controls RNAII transcripts that monitor the signaling mechanism through four proteins agrB, agrC, agrD, and agrA (Figure 1.8). AgrA-agrC complex forms the two component system where the sensor kinase agrC binds the extacellular AIP thereby activating agrA which in turn regulates P2 and P3 transcriptions. Thus, the level of the *agr* expression is controlled by changes in microenvironment depending on the bacterial growth rate and phase (Figure 1.9). Besides phase- or density-dependent regulation, a second level of global regulation is exerted by signals of environmental factors such as oxygen, nutrients, salt, osmolarity and pH which constitute the intracellular milieu. These induce control of the TCRS through a DNA-binding family of proteins called SarA and their homologues, as well as the sigmaB factor (Chan et al., 1998) leading to modulation of *agr* as well as downstream target genes (e.g., *hla* and *spa*) (Cheung et al., 1997; 2008) (Figure 1.8). *sar*A has three promoters P1, P3, and P2, respectively, the former is the predominant and the latter two are weaker, also P1 and P2 are σ^A dependent whereas, P3 is σ^B dependent and hence active at stationary or during stress times (Manna et al., 1998). As reviewed below, earlier studies using either direct assessment of regulator activities, or indirectly by monitoring expression of a specific gene (e.g., *clfA*) provided clues on how the bacterial microenvironment (organ, tissue, and intracellular milieu) regulates gene expression leading to evolution of adapted lineages.

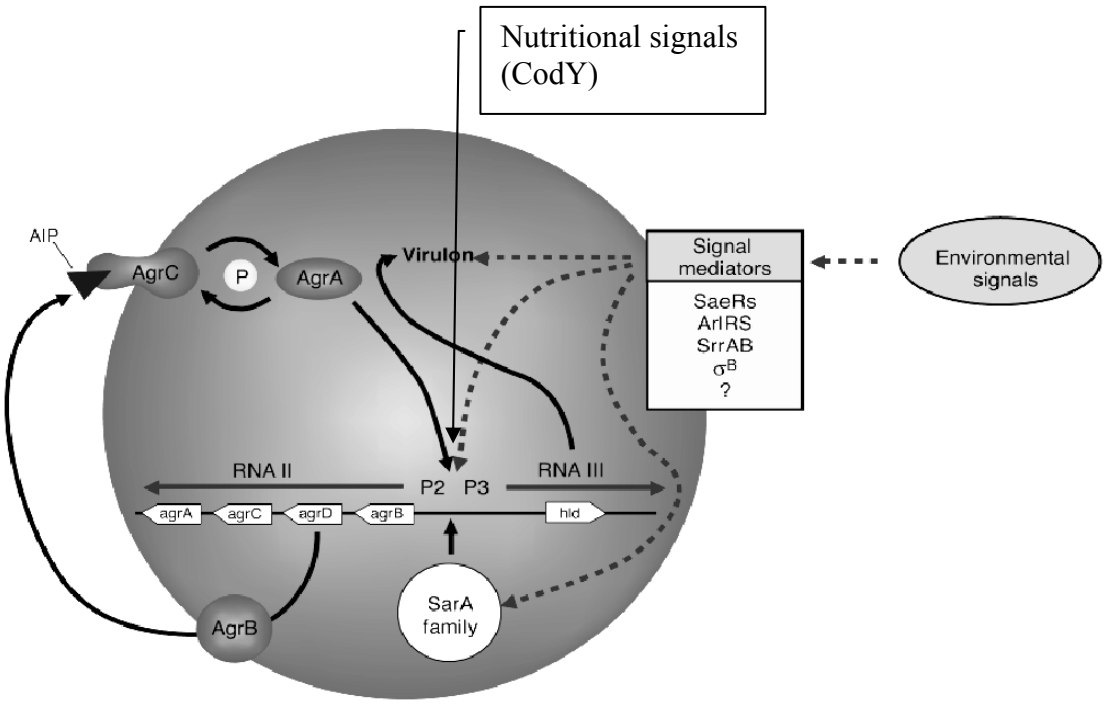


Figure 1.8. The accessory gene regulator (*agr*) system in *Staphylococcus*. The P2 operon encodes (via RNAII) the signaling mechanism, whereas the transcript of the P3 operon, RNAIII, acts as the effector molecule of the *agr*.

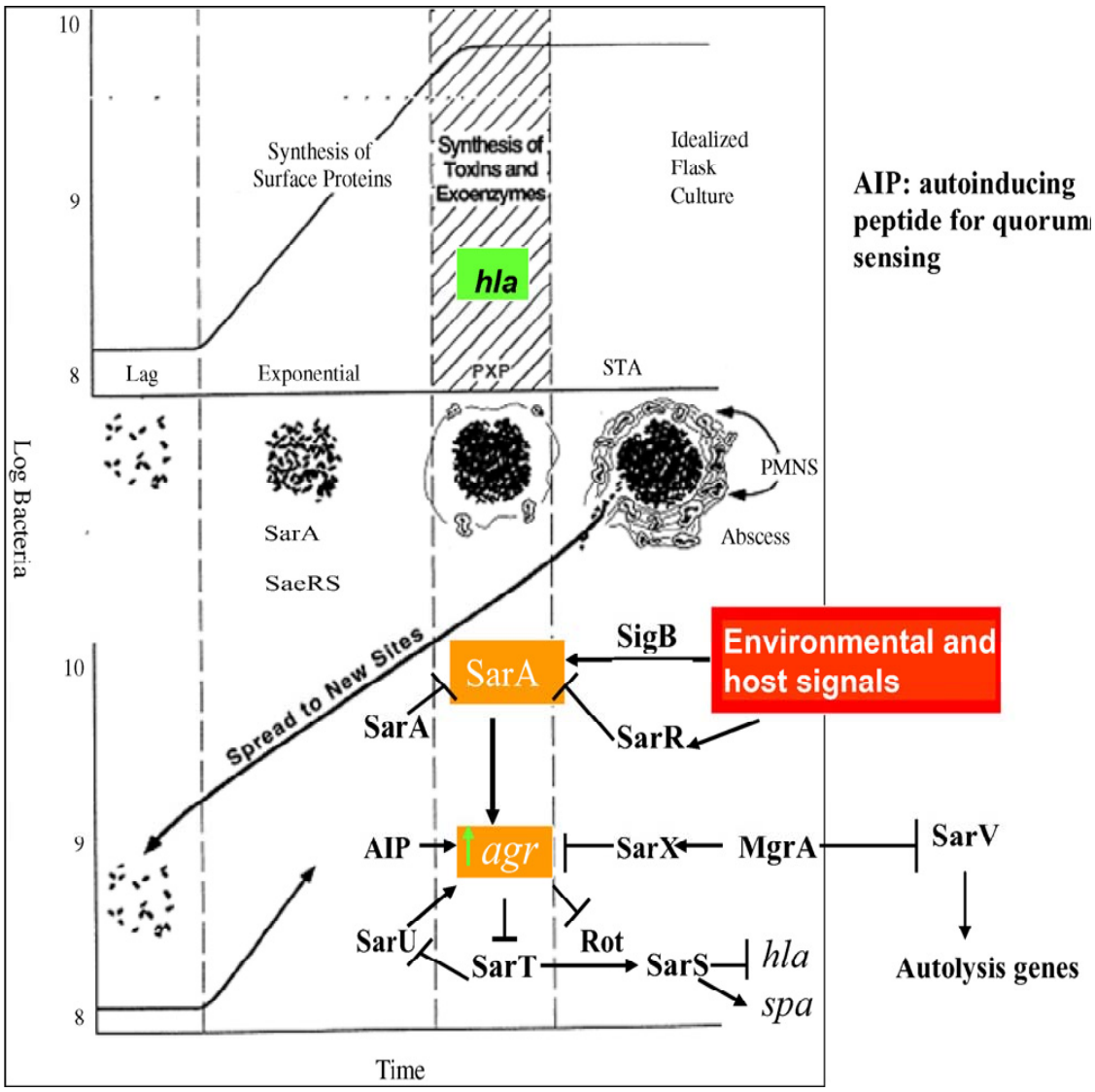


Figure 1.9. Regulation of virulence determinants in *S. aureus* by the SarA protein family (Cheung et al. 2008). Expression of surface adhesins coincide with the expression of SarA and SaeRS. From exponential to post exponential phase, production of extracellular toxins such as alpha-toxin would begin corresponding to the maximal expression of SarA and the ensuing activation of *agr*. SarA expression is repressed by SarA and SarR. SigB activates one of the *sarA* promoters (the P3 promoter). On the other hand, *agr* is controlled by SarA, a quorum sensing autoinducing peptide, other TCRS (see, MgrA, SarX and SarU). Activation of *agr* would lead to up-regulation of another TCRS system called SaeRS and down-regulation of a SarA protein homolog called Rot. This will eventually lead to repression of two gene products called SarT and subsequently SarS. SarT is an activator of SarS, which is a repressor of alpha toxin production and an activator of protein A synthesis, thus explaining the elevated production of alpha-toxin and repression of

protein A upon *agr* activation. Activation of *agr* would also result in the amplification of the original signal by activating SarU, which is a positive regulator of *agr*.

There are several lines of evidence to suggest that cells of specific organs play important roles in the regulation of virulence determinants and evolution of clones. *In vitro* gene fusion studies using laboratory strains revealed a role for SarA in signal transduction in response to microaerobiosis since levels of *hla* and toxins were found diminished in that condition (Chan and Foster, 1998). Whether this was due to a reduced quorum-sensing because of decreased growth yield, or related to an alternative unidentified regulatory pathway, could not be specified. Similarly, glucose and pH appeared to down regulate *agr* loci, perhaps due to catabolite repression of this locus (Regassa et al., 1992). Using the *S. aureus* laboratory strain RN6390 carrying a green fluorescent reporter system with *sarA* promoters P1, P2, and P3, the SarA *in vivo* activity was monitored during rabbit endocarditis (Cheung et al., 1998). At 24 hours the P1 was active *in vivo* and *in vitro* while P3 was inactive in both; whereas, P2 was active only *in vivo*. Most importantly, these studies showed that under same stable experimental conditions and animal model, promoter activities differed not only in different organs but also at different sites within the same organ. The promoter (*sarA* P2) was active in heart valve tissue, but not in kidney despite the prevalence of bacteria in both. It was also active on the surface of cardiac vegetation where the cells were metabolically active but not in the depth of the vegetation where the bacterial cells were dormant (Cheung et al., 1998). Furthermore, the host cell health state greatly influences the regulation of invasiveness. Strain RN6390 was able to aggressively enter, replicate, and escape in a cystic fibrosis cell line, but not in a normal lung epithelial cell culture where it was only able to enter and persist (Jarry et al., 2006). Thus, the ability to invade and persist within host cells could be specific to certain clones of *S. aureus*, and is the basis for the long-term chronic mastitis infections in dairy cattle.

Similarly, indirect assessment of regulator activities based on the expression patterns of a universal adhesion gene (e.g., *clfA*) indicates potential alternative regulatory pathways. The postulated ClfA activity early in infection (Wolz et al., 2002), although not consistent with the widely reported late transcription during growth *in vitro* (Wolz et al., 1996; Ni et al., 1998), correlates well with respect to alternative global regulation during

invasion. For instance, during the course of sepsis/septic arthritis in mice, expression of *clfA* was found to decrease over time in infected kidneys (representative of systemic infections), but increasing during the course of infection in joints (representing deep localized microenvironments). The authors suggested that specific tissue is decisive for the differentiation of staphylococci (Josefsson et al., 2008). Consistent with this, an SCV strain expressing fibronectin binding proteins and ClfA independently of the *agr* was more efficiently internalized by embryonic cells than were their isogenic controls (Vaudaux et al., 2002).

The bovine mammary gland is a complex environment and it is likely that unique expression profiles do exist. A recent study was conducted under low oxygenation condition mimicking the bovine mammary gland in order to monitor activities of a limited number of regulators and virulence factors, including *clfA*, by RealTime PCR. Irrespective of oxygen level and strain, the regulatory factors RNAIII, *rot*, and *sarR* genes, and the virulence factors asp23 and *clfA* genes were strongly expressed. Under low oxygenation, an enhanced expression of *srr*, *clfA* and *spa* genes was shown. However, some regulators such as *sae, sarA* and *sigB* were differentially transcribed according to the strain and the oxygenation condition (Ster et al., 2005). The authors suggested coexistence of alternative expression pathways for different strains and oxygen levels. Regulation of invasion into bovine mammary epithelial cell line (Bayles et al., 1998) has been shown to be controlled by the Agr and Sar systems. Mutants of these regulators were internalized by the cultured mammary cells at levels reproducibly greater than the wild type, but failed to induce apoptosis, as it required metabolically active bacteria (Wesson et al., 1998). In addition, inactivation of *hemB* in *S. aureus* strain Newman resulted in inactivation of *agr*, activation of *sigB* and its dependent transcripts such as *sarC* and *clfA*, alteration of expression pattern of global regulators, namely, *arl*, *rot*, *sae*, *sarR*, *sarS*, *srr*, *svrA*, and resulted in SCV phenotype (Senn et al., 2005). Thus, it is not clear how expression profiles change in the mammary gland.

Although studies on single/double mutant laboratory strains under a defined condition provides clues, there is a recent realization that bacterial colonization, invasion, and persistence or epidemicity genes operate as subsets under the control of complex regulatory network. In addition, many of the above experiments use long-term laboratory

passed strains that are prone to mutations, even though the disease might be caused by a well adapted host- or –tissue specific lineage. It has been shown that *in vitro* serial passage of *S. aureus* induces significant changes in physiology, virulence factor production, and more importantly, in *agr* nucleotide sequence (Somerville et al., 2002). For instance, the most commonly used strains 8325-4 (and its RN-derivatives) and strain Newman contain such mutations. The former strain has a known defect in *rsbU,* an anti-Sigma factor gene (Wertheim et al., 2008), and strain Newman has a frame-shift leading to stop codon that truncates fibronectin-binding proteins (Grundmeier et al., 2004). Thus, it is not known how gene expression profiles change during the mammary cell-*S.aureus* interactions.

S. aureus undergoes significant changes, in gene expression and physiology (Williams et al., 1999; Proctor et al., 2006) that are specific to host tissue. Thus, use of target cells might be more adequate. We hypothesize that bovine mammary cells might induce regulation of expression patterns unique to mammary gland. The genes expressed in this way during the infection process are mostly virulence genes that can be detected by expression profiling as distinct clusters. This is because, since virulence genes tend to be coordinated, they are also likely to be co-regulated. A justification (i.e., "shared expression implies shared function") was demonstrated in a system of cluster analysis for genome-wide expression data using *Saccharomyces cerevisiae* (Basset et al., 1999; Eisen et al., 1998). This approach has also been successful in identifying five clusters, including a specific subset of adherence genes, regulated upon contact of *Neisseria meningitides* serogroup B to human epithelial cells (Grifantini et al., 2002). Consequently, this approach was applied to identify phase-specific genes of *Neisseria meningitidis* serogroup B during different stages of infection, namely interactions with epithelial cells, endothelial cells, and serum (Kurz et al., 2003).

In summary, despite genome conservation, the mechanisms by which *S. aureus* adapt, not only to different hosts but also to different organs and tissues of the same host, and re-emerge as a virulent pathogen in healthy humans and dairy animals has been quite elusive. Genomic methods revealed the population genetic structure of the species, and typing by major molecular methods such as MLST defined the ancestral lines of the dominant clones successfully. However, these methods are not suitable for subtyping and

detecting evolution of subclonal populations that descend from dominant lines at different geographic locations and hosts. It is known that mastitis is caused by a limited number of global clonal lines; nevertheless, the distribution of dominant subtypes at different regions is still unknown. There is, thus, there is an urgent need for evaluating and developing subtyping methods that can couple the genotype diversity with their epidemicity and virulence phenotype in different hosts/organ and geographic locations. As reviewed above, since *S. aureus* responds to the microenvironment of host cells and differentiate into variant types, and the repeat containing surface proteins mediate staphylococcal speciation and host-pathogen interactions through ligand-binding and environmental signals transduction, we hypothesize that the coding intragenic repeats in surface protein of the clumping factor A, are one of factors responsible for adaptations and can be used to group strains with similar properties and to study the basic mechanisms underlying specialization in the mammary gland as well as different organs. This marker (*clfA* typing) would reveal the dominant subtypes with similar genotypic properties in different hosts/organs and regions. In addition, whole genome expression patterns are useful to study the mechanisms for invasion into mammary cells.

Working hypothesis and research objectives:

1. Working hypothesis

- *clfA* could be suitable for typing as well as studying host- and/or tissue-specialization in *S. aureus* lineages, due to its repeat stability, structure, and variability in *clfA*, along with the universality of the protein, its adaptive expression *in vivo*, and its direct involvement in virulence.

- *S. aureus* strains would have a tendency to adapt to a certain *clfA* length range to colonize and infect different organs in human (such as skin and upper respiratory tract) and in bovine such as the mammary gland. Thus, testing of clinical isolates from these sites would reveal organ specificity in the length of *clfA*. This hypothesis is based on the fact that the repeat region is the only variable domain in *clfA* and its ligand-binding region is conserved. The tandem repeats are

hypervariable among strains of this species. This variability affects the length of whole protein and consequently play important roles in adaptation,

- Interaction of a mastitis-specific isolate with a bovine mammary epithelial cell line called MAC-T, would lead to identification of expression patterns of subsets of genes necessary for invasion into host cells. These patterns of expression can be detected by whole genome systems such as high-throughput qPCR and microarrays. Gene subsets can be grouped based on their expression profiles such as binding and transport or metabolic pathways.

2. Research objectives:

- To confirm the suitability of *clfA* for use in typing and grouping *S. aureus.*
- To examine the R domain copy numbers and sequence information in independently isolated clinical isolates from infected sites such as skin and upper respiratory tract, and from bovine IMI isolates, from hosts located at different geographic regions for organ specificity in the length of *clfA*.
- To determine the gene expression profiles of mastitis and human specific *S. aureus* isolates during their invasion into MAC-T mammary epithelial cell line under carbon dioxide and reduced oxygen conditions.

REFERENCES

Ahmed S, Meghji S, Williams R, Henderson B, Brock J, Nair S. 2001. *Staphylococcus aureus* fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria. Infect Immun 69: 2872-2877.

Allos B, Moore M, Griffin P, Tauxe R. 2004. Surveillance for sporadic foodborne disease in the 21st century: the FoodNet perspective. Clin Infect Dis 38 Suppl 3: S115-120.

Almeida, R. A., K. R. Matthews and S. P. Oliver. 1997. Eukaryotic and Prokaryotic Cell Functions Required for Invasion of Staphylococcus aureus into Bovine Mammary Epithelial Cells. J. Vet. Med. B. 44:139-145.

Atkins K, Burman J, Chamberlain E, Cooper J, Poutrel B, Bagby S, Jenkins A, Feil E, van den Elsen J. 2008. S. aureus IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. Mol Immunol 45: 1600-1611.

Bancroft EA. 2007. Antimicrobial Resistance: It's Not Just for Hospitals *JAMA*. 298:1803-1804.Bar D, Gröhn Y, Bennett G, González R, Hertl J, Schulte H, Tauer L, Welcome F, Schukken Y. 2008. Effects of repeated episodes of generic clinical mastitis on mortality and culling in dairy cows. J Dairy Sci 91: 2196-2204.

Barkema H, Schukken Y, Lam T, Beiboer M, Wilmink H, Benedictus G, Brand A. 1998. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. J Dairy Sci 81: 411-419.

Bassett DJ, Eisen M, Boguski M. 1999. Gene expression informatics--it's all in your mine. Nat Genet 21: 51-55.

Bayles K, Wesson C, Liou L, Fox L, Bohach G, Trumble W. 1998. Intracellular *Staphylococcus aureus* escapes the endosome and induces apoptosis in epithelial cells. Infect Immun 66: 336-342.

Ben Zakour N, Guinane C, Fitzgerald J. 2008a. Pathogenomics of the staphylococci: insights into niche adaptation and the emergence of new virulent strains. FEMS Microbiol Lett 289: 1-12.

Blevins JS, Beenken KE, Elasri MO, Hurlburt BK, Smeltzer MS. 2002. Strain-dependent differences in the regulatory roles of sarA and agr in *Staphylococcus aureus*. Infect Immun. 70:470–80.

Bramley J. 1992. Identifying Mastitis Problems and Strategies for Control. Pages pp. 5–14. Proceedings of the 31st National Mastitis Council. Arlington, VA,

Brody T, Yavatkar A, Lin Y, Ross J, Kuzin A, Kundu M, Fann Y, Odenwald W. 2008. Horizontal gene transfers link a human MRSA pathogen to contagious bovine mastitis bacteria. PLoS One 3: e3074.

Brouillette E, Lacasse P, Shkreta L, Belanger J, Grondin G, Diarra MS, Fournier S, and Talbot BG. 2002. Immunization against the clumping factor (ClfA) *Staphylococcus aureus.* Vaccine 20:2348-2357.

Bronner S, Monteil H, Prévost G. 2004. Regulation of virulence determinants in *Staphylococcus aureus*: complexity and applications. FEMS Microbiol Rev 28:183-200.

Bubeck Wardenburg J, Palazzolo-Ballance A, Otto M, Schneewind O, DeLeo F. 2008. Panton-Valentine leukocidin is not a virulence determinant in murine models of community-associated methicillin-resistant *Staphylococcus aureus* disease. J Infect Dis 198: 1166-1170.

Burvenich C, Van Merris V, Mehrzad J, Diez-Fraile A, Duchateau L. 2003. Severity of E. coli mastitis is mainly determined by cow factors. Vet Res 34: 521-564.

Carter EW and Kerr DE. 2003. Optimization of DNA-based vaccination in cows using green fluorescent protein and protein A as a prelude in immunization against staphylococcal mastitis, J Dairy Sci 86:1177–1186.

Chan P, Foster S. 1998. Role of SarA in virulence determinant production and environmental signal transduction in *Staphylococcus aureus*. J Bacteriol 180: 6232-6241.

Chan P, Foster S, Ingham E, Clements M. 1998. The *Staphylococcus aureus* alternative sigma factor sigmaB controls the environmental stress response but not starvation survival or pathogenicity in a mouse abscess model. J Bacteriol 180: 6082-6089.

Cheng A, Kim H, Burts M, Krausz T, Schneewind O, Missiakas D. 2009. Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. FASEB J. 23:3393-404.

Cheung A, Bayer M, Heinrichs J. 1997. sar Genetic determinants necessary for transcription of RNAII and RNAIII in the agr locus of *Staphylococcus aureus*. J Bacteriol 179: 3963-3971.

Cheung A, Nast C, Bayer A. 1998. Selective activation of sar promoters with the use of green fluorescent protein transcriptional fusions as the detection system in the rabbit endocarditis model. Infect Immun 66: 5988-5993.

Cheung A, Nishina K, Trotonda M, Tamber S. 2008. The SarA protein family of *Staphylococcus aureus*. Int J Biochem Cell Biol 40: 355-361.

Carter EW and Kerr DE. 2003. Optimization of DNA-based vaccination in cows using green fluorescent protein and protein A as a prelude in immunization against staphylococcal mastitis, J Dairy Sci 86:1177–1186.

Citti C, Kim M, Wise K. 1997. Elongated versions of Vlp surface lipoproteins protect Mycoplasma hyorhinis escape variants from growth-inhibiting host antibodies. Infect Immun 65: 1773-1785.

Cossart P, Sansonetti PJ. 2004. Bacterial invasion: the paradigms of enteroinvasive pathogens. Science *304:*242-8.

Denis M, Wedlock D, Lacy-Hulbert S, Hillerton J, Buddle B. 2009. Vaccines against bovine mastitis in the New Zealand context: what is the best way forward? N Z Vet J 57: 132-140.

Dordet-Frisoni E, Dorchies G, De Araujo C, Talon R, Leroy S. 2007. Genomic diversity in *Staphylococcus xylosus*. Appl Environ Microbiol 73: 7199-7209.

Dosogne H, Vangroenweghe F, Burvenich C. 2002. Potential mechanism of action of J5 vaccine in protection against severe bovine coliform mastitis. Vet Res 33: 1-12.

Eisen M, Spellman P, Brown P, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863-14868.

Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. 2002.The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci U S A 99:7687-7692.

Epstein JH and Price JT. 2009. The significant but understudied impact of pathogen transmission from humans to animals. Mt Sinai J Med 76:448-455.

Ericsson Unnerstad H, Lindberg A, Persson Waller K, Ekman T, Artursson K, Nilsson-Ost M, Bengtsson B. 2009. Microbial aetiology of acute clinical mastitis and agent-specific risk factors. Vet Microbiol 137: 90-97.

Erskine R, Eberhart R, Hutchinson L, Spencer S, Campbell M. 1988. Incidence and types of clinical mastitis in dairy herds with high and low somatic cell counts. J Am Vet Med Assoc 192: 761-765.

Feil EJ, Cooper JE, Grundmann H, Robinson DA,. Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt B G, Moore C E, Day NP J..2003. How clonal is *Staphylococcus aureus*? J Bacteriol 185: 3307-3316.

Feil, E, Li B, Aanensen D, Hanage W, Spratt B. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186**:**1518-30.

Finlay, BB, Cossart P. 1997. Exploitation of mammalian host cell functions by bacterial pathogens. Science, 276 : 718-725.

Fitzgerald J, Meaney W, Hartigan P, Smyth C, Kapur V. 1997. Fine-structure molecular epidemiological analysis of *Staphylococcus aureus* recovered from cows. Epidemiol Infect 119: 261-269.

Foster T. 2005. Immune evasion by staphylococci. Nat Rev Microbiol 3: 948-958.

Foster T, Höök M. 1998. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 6: 484-488.

Goff J. 2006. Major advances in our understanding of nutritional influences on bovine health. J Dairy Sci 89: 1292-1301.

Gravekamp C, Rosner B, Madoff L. 1998. Deletion of repeats in the alpha C protein enhances the pathogenicity of group B streptococci in immune mice. Infect Immun 66: 4347-4354.

Grifantini R, Bartolini E, Muzzi A, Draghi M, Frigimelica E, Berger J, Ratti G, Petracca R, Galli G, Agnusdei M, Giuliani MM, Santini L, Brunelli B, Tettelin H, Rappuoli R, Randazzo F, Grandi G.2002. Previously unrecognized vaccine candidates against group B meningococcus identified by DNA microarrays. Nat Biotechnol 20: 914-921.

Grundmeier M, Hussain M, Becker P, Heilmann C, Peters G, Sinha B. 2004. Truncation of fibronectin-binding proteins in *Staphylococcus aureus* strain Newman leads to deficient adherence and host cell invasion due to loss of the cell wall anchor function. Infect Immun 72: 7155-7163.

Grohn Y, Wilson D, González R, Hertl J, Schulte H, Bennett G, Schukken Y. 2004. Effect of pathogen-specific clinical mastitis on milk yield in dairy cows. J Dairy Sci 87: 3358-3374.

Guinane C, Sturdevant D, Herron-Olson L, Otto M, Smyth D, Villaruz A, Kapur V, Hartigan P, Smyth C, Fitzgerald J. 2008. Pathogenomic analysis of the common bovine *Staphylococcus aureus* clone (ET3): emergence of a virulent subtype with potential risk to public health. J Infect Dis 197: 205-213.

Hartford O, Francois P, Vaudaux P, Foster T. 1997. The dipeptide repeat region of the fibrinogen-binding protein (clumping factor) is required for functional expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. Mol Microbiol 25: 1065-1076.

Hidron AI, Low CE, Honig EG, Blumberg HM.2009. Emergence of community-acquired meticillin-resistant *Staphylococcus aureus* strain USA300 as a cause of necrotising community-onset pneumonia. Lancet Infect Dis. 9:384-392.

Higgins J, Loughman A, van Kessel K, van Strijp J, Foster T. 2006. Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. FEMS Microbiol Lett 258: 290-296.

Highlander SK, Hultén KG, Qin X, Jiang H, Yerrapragada S, Mason EO Jr, Shang Y, Williams TM, Fortunov RM, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox GE, Cardenas AM, Muzny DM, Hemphill L, Ding Y, Dugan S, Blyth PR, Buhay CJ, Dinh HH, Hawes AC, Holder M, Kovar CL, Lee SL, Liu W, Nazareth LV, Wang Q, Zhou J, Kaplan SL, Weinstock GM.2007. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. BMC Microbiol 7: 99.

Jarry T, Cheung A. 2006. *Staphylococcus aureus* escapes more efficiently from the phagosome of a cystic fibrosis bronchial epithelial cell line than from its normal counterpart. Infect Immun 74: 2568-2577.

Jernigan JA, Pullen AL, Partin C, Jarvis WR. 2003. Prevalence of and risk factors for colonization with methicillin-resistant *Staphylococcus aureus* in an outpatient clinic population. Infect Control Hosp Epidemiol. 24:445–450.

Jordan P, Snyder L, Saunders N. 2003. Diversity in coding tandem repeats in related Neisseria spp. BMC Microbiol 3: 23.

Josefsson E, Kubica M, Mydel P, Potempa J, Tarkowski A. 2008. In vivo sortase A and clumping factor A mRNA expression during *Staphylococcus aureus* infection. Microb Pathog 44: 103-110.

Josefsson E, McCrea K, Ní Eidhin D, O'Connell D, Cox J, Höök M, Foster T. 1998. Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus*. Microbiology 144 (Pt 12): 3387-3395.

Juhász-Kaszanyitzky E, Jánosi S, Somogyi P, Dán A, van der Graaf-van Bloois L, van Duijkeren E, Wagenaar J. 2007. MRSA transmission between cows and humans. Emerg Infect Dis 13: 630-632.

Kapur V, Sischo W, Greer R, Whittam T, Musser J. 1995. Molecular population genetic analysis of *Staphylococcus aureus* recovered from cows. J Clin Microbiol 33: 376-380.

Keefe G, Dohoo I, Spangler E. 1997. Herd prevalence and incidence of Streptococcus agalactiae in the dairy industry of Prince Edward Island. J Dairy Sci 80: 464-470.

Kennedy AD, Otto M, Braughton KR, Whitney A R, Chen L, Mathema B, Mediavilla JR, Byrne KA, Parkins LD, Tenover F C, Kreiswirth BN, Musser JM, and DeLeo F R.2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. Proc Natl Acad Sci U S A 105: 1327-1332.

Klevens R, M. A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L.H. Harrison, R.Lynfield, G. Dumyati, J.M. Townes, A. S. Craig, E.R. Zell, G.E. Fosheim, L.K. McDougal, R. B. Carey, S.K. Fridkin, for the Active Bacterial Coresurveillance (ABCs) MRSA Investigators. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. JAMA 298: 1763-1771.

Kluytmans J, van Belkum A, Verbrugh H. 1997. Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev 10:505–520.

Koreen L, Ramaswamy S, Graviss E, Naidich S, Musser J, Kreiswirth B. 2004. spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol 42**:**792-9.

Kuhn G, Francioli P, Blanc D. 2006. Evidence for clonal evolution among highly polymorphic genes in methicillin-resistant *Staphylococcus aureus*. J Bacteriol 188: 169-178.

Kurz S, Hübner C, Aepinus C, Theiss S, Guckenberger M, Panzner U, Weber J, Frosch M, Dietrich G. 2003. Transcriptome-based antigen identification for Neisseria meningitidis. Vaccine 21: 768-775.

Lammerding A. 1998. An overview of microbial food safety risk assessment. J Food Prot 60: 1420-1425.

LeBlanc S, Lissemore K, Kelton D, Duffield T, Leslie K. 2006. Major advances in disease prevention in dairy cattle. J Dairy Sci 89: 1267-1279.

Lee J. 2003. Methicillin (Oxacillin)-resistant *Staphylococcus aureus* strains isolated from major food animals and their potential transmission to humans. Appl Environ Microbiol 69: 6489-6494.

Li M, Diep B, Villaruz A, Braughton K, Jiang X, DeLeo F, Chambers H, Lu Y, Otto M. 2009. Evolution of virulence in epidemic community-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci U S A 106: 5883-5888.

Lindsay J, Moore C, Day N, Peacock S, Witney A, Stabler R, Husain S, Butcher P, Hinds J. 2006. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J Bacteriol 188: 669-676.

Madoff L, Michel J, Gong E, Kling D, Kasper D. 1996. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc Natl Acad Sci U S A 93: 4131-4136.

Manna A, Bayer M, Cheung A. 1998. Transcriptional analysis of different promoters in the sar locus in *Staphylococcus aureus*. J Bacteriol 180: 3828-3836.

Mathew A, Cissell R, Liamthong S. 2007. Antibiotic resistance in bacteria associated with food animals: a United States perspective of livestock production. Foodborne Pathog Dis 4: 115-133.

McAleese F, Walsh E, Sieprawska M, Potempa J, Foster T. 2001. Loss of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation of transcription, shedding and cleavage by metalloprotease. J Biol Chem 276: 29969-29978.

McDougall S. 1998. Efficacy of two antibiotic treatments in curing clinical and subclinical mastitis in lactating dairy cows. N Z Vet J 46: 226-232.

Melles D, Gorkink R, Boelens H, Snijders S, Peeters J, Moorhouse M, van der Spek P, van Leeuwen W, Simons G, Verbrugh H, van Belkum A. 2004. Natural population dynamics and expansion of pathogenic clones of *Staphylococcus aureus*. J Clin Invest 114**:**1732-40.

Michie CA. 2002. Staphylococcal vaccines. Trends in Immunology 23: 461-463.

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

Novick, RP. 2003. Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol 48:1429–1449.

Olde Riekerink R, Barkema H, Kelton D, Scholl D. 2008. Incidence rate of clinical mastitis on Canadian dairy farms. J Dairy Sci 91: 1366-1377.

Palmqvist N, Josefsson E, Tarkowski A. 2004a. Clumping factor A-mediated virulence during *Staphylococcus aureus* infection is retained despite fibrinogen depletion. Microbes Infect 6: 196-201.

Palmqvist N, Patti J, Tarkowski A, Josefsson E. 2004b. Expression of staphylococcal clumping factor A impedes macrophage phagocytosis. Microbes Infect 6: 188-195.

Proctor R, von Eiff C, Kahl B, Becker K, McNamara P, Herrmann M, Peters G. 2006. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. Nat Rev Microbiol 4: 295-305.

Projan S, Novick R.1997. The molecular basis of pathogenicity. In: Archer G,Crossley KStaphylococci in human diseases. New York, N.Y.: Churchill Livingstone pp. 55-81.

Regassa L, Novick R, Betley M. 1992. Glucose and nonmaintained pH decrease expression of the accessory gene regulator (agr) in *Staphylococcus aureus*. Infect Immun 60: 3381-3388.

Rabello RF, Moreira BM, Lopes RM, Teixeira LM, Riley LW, Castro AC.2007. Multilocus sequence typing of *Staphylococcus aureus* isolates recovered from cows with mastitis in Brazilian dairy herds. J Med Microbiol 56:1505-1511.

Reksen O, Sølverød L, Branscum A, Osterås O. 2006. Relationships between milk culture results and treatment for clinical mastitis or culling in Norwegian dairy cattle. J Dairy Sci 89: 2928-2937.

Risley A, Loughman A, Cywes-Bentley C, Foster T, Lee J. 2007. Capsular polysaccharide masks clumping factor A-mediated adherence of *Staphylococcus aureus* to fibrinogen and platelets. J Infect Dis 196: 919-927.

Roberson JR, Fox LK, Hancock DD, Gay JM, Besser TE. 1998. Sources of intramammary infections from *Staphylococcus aureus* in dairy heifers at first parturition. J Dairy Sci 81:687–693.

Roberson J, Warnick L, Moore G. 2004. Mild to moderate clinical mastitis: efficacy of intramammary amoxicillin, frequent milk-out, a combined intramammary amoxicillin, and frequent milk-out treatment versus no treatment. J Dairy Sci 87: 583-592.

Robinson D, and Enright M. 2004. Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. Clin Microbiol Infect 10:92-7.

Rossney A, Shore A, Morgan P, Fitzgibbon M, O'Connell B, Coleman D. 2007. The emergence and importation of diverse genotypes of methicillin-resistant *Staphylococcus aureus* (MRSA) harboring the Panton-Valentine leukocidin gene (pvl) reveal that pvl is a poor marker for community-acquired MRSA strains in Ireland. J Clin Microbiol 45: 2554-2563.

Sabat A, Krzyszton-Russjan J, Strzalka W, Filipek R, Kosowska K, Hryniewicz W, Travis J, Potempa J. 2003. New method for typing *Staphylococcus aureus* strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates. J Clin Microbiol 41:1801-1804.

Salisbury, JG, Nicholl TJ, Lammerding AM, Turnidge J, and Nunn MJ. 2002. A risk analysis framework for the long-term management of antibiotic resistance in food-producing animals. *Int. J. Antimicrob. Agts*. 20: 153-164.

Schneewind O, Fowler A, Faull K. 1995. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. Science 268: 103-106.

Seegers H, Fourichon C, Beaudeau F. Production effects related to mastitis and mastitis economics in dairy cattle herds. Vet Res 34: 475-491.

Senn M, Bischoff M, von Eiff C, Berger-Bächi B. 2005. sigmaB activity in a *Staphylococcus aureus* hemB mutant. J Bacteriol 187: 7397-7406.

Shompole S, Kim TH, Linda E L, Katarzyna D, Gregory AB and Kenneth WB. 2003.Biphasic intracellular expression of *Staphylococcus aureus* virulence factors and evidence for Agr-mediated diffusion sensing. Mol. Microbiol. 49:919–927.

Simmons W, Denison A, Dybvig K. 2004. Resistance of Mycoplasma pulmonis to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. Infect Immun 72: 6846-6851.

Sinha B, François PP, Nüsse O, Foti M, Hartford OM, Vaudaux P, Foster TJ, Lew DP, Herrmann M, Krause KH. 1999**.** Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin α5ß1. Cell Microbiol. **1**:101-1**1**7.

Sivaraman K, Cole AM. 2009. Pathogenesis gene families in the common minimal genome of *Staphylococcus aureus* are hypervariable. FEBS Lett 583:1304-1308

Smith E, Green L, Medley G, Bird H, Fox L, Schukken Y, Kruze J, Bradley A, Zadoks R, Dowson C. 2005. Multilocus sequence typing of intercontinental bovine *Staphylococcus aureus* isolates. J Clin Microbiol 43: 4737-4743.

Somerville G, Beres S, Fitzgerald J, DeLeo F, Cole R, Hoff J, Musser J. 2002. In vitro serial passage of *Staphylococcus aureus*: changes in physiology, virulence factor production, and agr nucleotide sequence. J Bacteriol 184: 1430-1437.

Stephen C, Artsob H, Bowie WR, Drebot M, Fraser E, Leighton T, Morshed M, Ong C, Patrick D. 2004. Prespectives on ermging zoonotic disease research and capacity building in Canada. Can J Infect Dis Med Microbiol 15:339-344.

Ster C, Gilbert F, Cochard T, Poutrel B. 2005. Transcriptional profiles of regulatory and virulence factors of *Staphylococcus aureus* of bovine origin: oxygen impact and strain-to-strain variations. Mol Cell Probes 19: 227-235.

Struelens MJ, Hawkey PM, French GL, Witte W, Tacconelli E. 2009. Laboratory tools and strategies for methicillin-resistant *Staphylococcus aureus* screening, surveillance and typing: state of the art and unmet needs. Clin Microbiol Infect. 15:112-119.

Sung J, Lindsay J. 2007. *Staphylococcus aureus* strains that are hypersusceptible to resistance gene transfer from enterococci. Antimicrob Agents Chemother 51: 2189-2191.

Sweet RL, Gibbs RS. Clinical microbiology of the female genital tract. 2001. In: Sweet RL, Gibbs RS, editors. *Infectious Diseases of the Female Genital Tract.* 4th. Philadelphia, Pa, USA: Lippincott Williams & Wilkins;. pp. 3–12.

Takeuchi F, Takeuchi F, Watanabe S, Baba T, Yuzawa H, Ito T, Morimoto Y, Kuroda M, Cui L, Takahashi M, Ankai A, Baba S, Fukui S, Lee JC, Hiramatsu K. 2005. Whole-genome sequencing of *staphylococcus haemolyticus* uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. J Bacteriol 187: 7292-7308.

Todd WJ. 2005. Community-Associated Methicillin-Resistant *Staphylococcus aureus.* CID 41:S269

Tuchscherr L, Buzzola F, Alvarez L, Caccuri R, Lee J, Sordelli D. 2005. Capsule-negative *Staphylococcus aureus* induces chronic experimental mastitis in mice. Infect Immun 73: 7932-7937.

van Belkum A. 2007. Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). FEMS Immunol Med Microbiol 49: 22-27.

van Belkum A, Scherer S, van Alphen L, Verbrugh H. 1998a. Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev 62: 275-293.

van Belkum A, Hermans P, Licciardello L, Stefani S, Grubb W, van Leeuwen W, Goessens W. 1998b. Polymerase chain reaction-mediated typing of microorganisms: tracking dissemination of genes and genomes. Electrophoresis 19: 602-607.

van Belkum A, Scherer S, van Leeuwen W, Willemse D, van Alphen L, Verbrugh H. 1997. Variable number of tandem repeats in clinical strains of Haemophilus influenzae. Infect Immun 65: 5017-5027.

Vandenesch F, Naimi T, Enright MC, Lina G, Nimmo GR, Heffernan H, Liassine N, Bes M, Greenland T, Reverdy ME, Etienne J. 2003. Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. Emerg Infect Dis 9:978-984.

van Leeuwen W, Melles D, Alaidan A, Al-Ahdal M, Boelens H, Snijders S, Wertheim H, van Duijkeren E, Peeters J, van der Spek P, Gorkink R, Simons G, Verbrugh H, van Belkum A. 2005. Host- and tissue-specific pathogenic traits of *Staphylococcus aureus*. J Bacteriol 187**:**4584-4591.

van Loo I, Huijsdens X, Tiemersma E, de Neeling A, van de Sande-Bruinsma N, Beaujean D, Voss A, Kluytmans J. 2007. Emergence of methicillin-resistant *Staphylococcus aureus* of animal origin in humans. Emerg Infect Dis 13: 1834-1839.

Vaudaux P, Francois P, Bisognano C, Kelley W, Lew D, Schrenzel J, Proctor R, McNamara P, Peters G, Von Eiff C. 2002. Increased expression of clumping factor and fibronectin-binding proteins by hemB mutants of *Staphylococcus aureus* expressing small colony variant phenotypes. Infect Immun 70: 5428-5437.

Waldron D, Lindsay J. 2006. Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between S. aureus isolates of different lineages. J Bacteriol 188: 5578-5585.

Wagner, G., L. Bohr, P. Wagner, and L. N. Petersen. 1984. Tampon-induced changes in vaginal oxygen and carbon dioxide tensions. Am. J. Obstet. Gynecol. 148:147-150

Wertheim H, Walsh E, Choudhurry R, Melles D, Boelens H, Miajlovic H, Verbrugh H, Foster T, van Belkum A. 2008. Key role for clumping factor B in Staphylococcus aureus nasal colonization of humans. PLoS Med 5: e17.

Weese JS. 2005. Methicillin-Resistant Staphylococcus aureus: An Emerging Pathogen in Small Animals. J Am Anim Hosp Assoc 41:150-157.

Weese JS, Rousseau J, Traub-Dargatz JL, Willey BM, McGeer AJ, E.Low D. 2005. Community-associated methicillin-resistant *Staphylococcus aureus* in horses and humans who work with horses. J Am Vet Med Assoc. 226:580-583.

Wesson C, Liou L, Todd K, Bohach G, Trumble W, Bayles K. 1998. *Staphylococcus aureus* Agr and Sar global regulators influence internalization and induction of apoptosis. Infect Immun 66: 5238-5243.

Williams I, Williams I, Paul F, Lloyd D, Jepras R, Critchley I, Newman M, Warrack J, Giokarini T, Hayes AJ, Randerson PF, Venables WA. 1999. Flow cytometry and other techniques show that *Staphylococcus aureus* undergoes significant physiological changes in the early stages of surface-attached culture. Microbiology 145: 1325-1333.

Wolz C, McDevitt D, Foster T, Cheung A. 1996. Influence of agr on fibrinogen binding in *Staphylococcus aureus* Newman. Infect Immun 64: 3142-3147.

Wolz C, Goerke C, Landmann R, Zimmerli W, Fluckiger U. 2002. Transcription of clumping factor A in attached and unattached *Staphylococcus aureus* in vitro and during device-related infection. Infect Immun 70: 2758-2762.

Yarwood JM, McCormick JK, Schlievert PM. 2001. Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of Staphylococcus aureus. J Bacteriol 183:1113–23.

Yarwood J, Schlievert P. 2003. Quorum sensing in Staphylococcus infections. J Clin Invest 112: 1620-1625.

Zhang K, McClure J, Elsayed S, Tan J, Conly J. 2008. Coexistence of Panton-Valentine leukocidin-positive and -negative community-associated methicillin-resistant *Staphylococcus aureus* USA400 sibling strains in a large Canadian health-care region. J Infect Dis 197: 195-204.

Zadoks RN, van Leeuwen WB, Kreft D, Fox LK, Barkema HW, Schukken YH, van Belkum A .2002. Comparison of *Staphylococcus aureus* isolates from bovine and human skin, milking equipment, and bovine milk by phage typing, pulsed-field gel electrophoresis, and binary typing. J Clin Microbiol 40:3894-3902.

Zecconi A, Piccinini R and Fox LK. 2003. Epidemiologic study of intramammary infections with *Staphylococcus aureus* during the program control in nine commercial dairy herds. J Am Vet Med Assoc 223:684–688.

CHAPTER II. REPEAT-BASED SUBTYPING AND GROUPING OF STAPHYLOCOCCUS AUREUS FROM HUMAN INFECTIONS AND BOVINE MASTITIS USING THE R-DOMAIN OF THE CLUMPING FACTOR A GENE

Kamaleldin B Said[1], Karam Ramotar[2,3] , Guoqiang Zhu[1,4], and Xin Zhao[1*]

[1]McGill University, Department of Animal Science, Ste. Anne de Bellevue, Quebec H9X 3V9, Canada

[2]Department of Pathology and Laboratory Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.

[3]Division of Microbiology, Ottawa Hospital, Ottawa, ON, Canada

[4] College of Veterinary Science, Yanzhou University, China

[*]Corresponding author. Tel.: 1-514 398 7975; Fax: 1-514 398 7964.

Diagnostic Microbiolology and Infectious Disease. 2009. 63:24-37.

ABSTRACT

Staphylococcus aureus has become an emerging public health concern. Markers capable of differentiating separate host-specific lineages are needed for tracing strain sources. Thus, a coding variable-number tandem-repeat (VNTR)-based typing was explored in this study, based on R-domain of clumping factor A (*clfA*) gene. DNA from isolates and strains of human infections and bovine mastitis were amplified and sequenced. Sequences of *clfA* from published strains were also analyzed. Results indicate that except one with 36 copies, 44 of the 55 R-domains had repeat-copies between 44 and 57, while the remaining ten had 59.5-73 copies. Furthermore, human isolates were polymorphic, while mastitis isolates were clonal. Phylogenetic grouping assigned host-specific strains into respective clusters. The repeats were stable during passages in milk, nutrient broth, and invasion of mammary cells showing suitability for typing. Our data show that the R-domain can be useful for typing and grouping host specific lineages. Moreover, existence of variant repeats in human strains and the dominance of a clonal motif in mastitis may imply that a specific selection has occurred in the mammary gland.

INTRODUCTION

Staphylococcus aureus is an increasing public health concern due to the rapid evolution and spread of virulent/resistant clones in humans and animals. It has been rated as the second in the four leading causes of nosocomial bloodstream infections, with a mortality rate of 25 percent (Wenzel and Edmond, 2001). In the dairy industry, bovine mastitis is the most frequent and most costly disease (Seegers et al., 2003) and *S. aureus* has been one of the most important causative agents of intramammary infections (Wilson et al., 1997).

Several typing methods have been used in classifying *S. aureus* strains from human infections and bovine mastitis, including phage typing, pulsed field gel electrophoresis (PFGE), binary typing, multilocus sequence typing (MLST), as well as phenotypic traits (Roberson et al., 1998; Zadoks et al., 2002; Enright et al., 2002; Rabello et al., 2007). PFGE has been the gold standard for typing. In addition, variable-number tandem-repeat (VNTR) typing has been increasingly suggested as a useful combination to major molecular methods such as PFGE and DNA-arrays. The first typical VNTR system for *S. aureus* was developed by Sabat et al., (2003) and used PCR amplicons of repeats from *sdr, clfA, clfB, ssp* and *spa* genes. When the VNTR system was compared with PFGE, it performed well; overall strain clustering was highly congruent and the type cut-off could be defined at 70–80%, equivalent to that of PFGE values of 75–90% (Malachowa et al., 2005). The *spa* typing has also been employed as a single genetic marker (Koreen et al., 2004). Like *spa* typing, typing by *clf*-R-domains has also been postulated as useful additions to major molecular methods; particularly for detecting microvariations in emerging clones in human infections (Gomes et al., 2005; Koreen et al., 2004, 2005). The R-domain of *clfA* is composed of 18bp repeats that follow the consensus [GAP (P= pyrimidine), TCN (N=any base), GAP, TCN, GAP, AGP]; they are in-frame for translation and therefore code for a $[Ser\text{-}Asp]_3$ dipeptide repeat (Foster and Hook, 1998). Repeat regions in *clfA, clfB, spa,* and fibronectin binding protein A gene (*fnbA*) have been applied as sequence-based typing on three sets of 48 methicillin-resistant *Staphylococcus aureus* (MRSA) strains by Kuhn et al., (2007). However, this

study was not considered VNTR-based since variant alleles were determined based on differences in nucleotide sequences and not from variations in repeat-copy numbers.

Intragenic repeats are potential systems for subtyping variant clones of clinical strains such as the small colony variants, community and hospital MRSAs, and the catalase negative MRSA, where major methods fail (Del'Alamo et al., 2007). Coding VNTRs are also promising systems for both basic and applied research based on many established findings. For instance, analysis of house-keeping as well as of more polymorphic genes in MRSA described *S. aureus* as having clonal evolution and genome (Feil et al., 2003; Kuhn et al., 2006). Similarly, comparative genomic analysis revealed that the strains causing infections in humans and their associated pets have the same genetic background and virulence gene content (van Leeuwen et al., 2005). Nevertheless, mastitis-associated *S. aureus* formed distinct clusters of tissue specific clones, indicating predilection to mammary glands (van Leeuwen et al., 2005). This is also consistent with the notion that specific tissue environment is decisive for the differentiation of staphylococci (Josefsson et al., 2008). In the light of the clonality, the genetic basis for the selective differentiation or host-specialization is currently ill defined. However, one potential tool for typing and studying host-specific clones would be tandem intragenic repeats. This is because pathogens tend to use coding repeats for long-term adaptation (Verstrepen et al., 2005). Similarly, it has been shown that copy number of repeats affect adherence and clumping titers of the cocci (Risley et al., 2007; Hartford et al., 1997).

Repeats chosen for typing should be reasonably stable. Some microsatellites and minisatellites are not suitable as markers for typing. For example, microsatellites known for immediate phase variations in pathogenic bacteria with dynamic genomes, such as *H. influenzae* and *N. meningitidis* (Bayliss et al. 2001; Moxon et al., 2006) are not suitable markers for typing. Similarly, some of longer coding repeats are used for immediate adaptation to host (Ayraud et al. 2003; Salaun et al., 2005) and such repeats are not useful for typing purposes. On the other hand, some minisatellites are reasonably stable and suitable for typing. The recent study by Kuhn et al., (2007) has indicated that *clfA* was much more stable than other markers in the clonal MRSA backgrounds. However, it would also be equally important to subject the bacteria for rigorous serial passaging under different media and conditions to see the stability of the repeat. Therefore, in this study,

we aimed to evaluate the potential use of *clfA* R-domain repeat-based system for typing diverse *S. aureus* strains and isolates from human infections and bovine mastitis as well as grouping host specific lineages.

MATERIALS AND METHODS

Bacterial strains

In this study, *clfA* sequences from 55 *S. aureus* strains/isolates from human infections and bovine mastitis were studied, including nine from completed *S. aureus* genomes sequences{ (representing hospital-acquired methicillin-resistant *S. aureus* (HA-MRSA), community-acquired MRSA (CA-MRSA), and vancomycin-resistant *S. aureus* (VRSA-MRSA) Table 2.1}, 24 MSSA (methicillin sensitive *S. aureus*) isolates from blood and synovial fluid (SF) of diverse populations of patients (from Ottawa Hospital, Ottawa, ON, Canada) and 19 bovine clinical mastitis isolates of *S. aureus* from different regions of Quebec dairy herds (by the clinical laboratory of Faculté de médecine vétérinarire, Université de Montreal, Quebec, Canada. These isolates were generously provided by Dr. Serge Messier. In addition, three reference *S. aureus* strains Lowenstein (ATCC49521), Wright (ATCC49525), and Smith Compact (noncapsulated variant BAA-934) were included in the study. Stocks were maintained frozen at –80°C in trypticase soy broth (BD, Sparks, MD, USA) before experiments. Both bovine mastitis isolates and those from human infections were MSSAs.

Table 2.1. Names and addresses of published *S. aureus* genomes used in the analyses of R-domain copy numbers of *clfA*.

Strain	**Source**	**Site**	**Origin**	**Copy #**
S.aureus aureus COL (MRSA)	TIGR	http://www.tigr.org	Human	50.5
S.aureus aureus MRSA252 (HA-MRSA= hospital acquired methicillin resistant *S. aureus*)	Welcome Trust Sanger Institute	http://www.sanger.ac.uk/	Human	66
S.aureus aureus N315 HA-MRSA	Juntendo Univ.	http://www.staphylococcus.org/jp/	Human	59.5
S. aureus aureus MSSA476 (CA-MSSA=community acquired methicillin sensitive *S. aureus*) hyper-virulent	Welcome Trust Sanger Institute	http://www.sanger.ac.uk/	Human	49
S.aureus aureus MW2 (CA-MRSA)	NITE	http://www.bio.nite.go.jp/	Human	51
S.aureus Mu50 MRSA-VRSA (MRSA-Vancomycin resistant *S. aureus*)	Juntendo Univ.	http://www.staphylococcus.org/jp/	Human	49.5
Staphylococcus aureus NCTC 8325	University of Oklahoma	http://microgen.ouhsc.edu/s_aureus/s_aureus_home.htm	Human	48
Staphylococcus aureus USA300 strong association with unusually invasive disease (invasive-MRSA)	University of California, San Francisco	http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj&cmd=Retrieve&dopt=Overview&list_uids=16313	Human	49
Mastitis *S. .aureus RF122*	Univ. Minnesota	http://pathogenomics.umn.edu/staph_index.htm	Bovine mastitis	46

Preparation of milk samples

Preparation of normal fresh bovine milk was carried out as described by Barrio et al., (2003). Fresh bovine milk samples were obtained aseptically from a healthy cow housed on the Macdonald Campus Farm, McGill University. Samples were centrifuged at 1000 g for 15 min at 4 ºC to remove cells and most of the cream and the resulting sample was termed as normal milk. A portion of normal milk was heated at 80 ºC for one hour (Lammers et al., 2000) for sterilizing and was termed as heat-sterilized milk. In addition, commercial micromembrane filtered milk was purchased from a local supermarket (Lactantia Purfiltre, Parmalate Canada) and was termed as commercial milk. All milk samples were determined free of bacteria upon bacteriological analysis.

Pilot experiment

A pilot experiment was conducted during a 3-day passage of S. *aureus* Wright to test whether different sizes of bacterial inocula and DNA extraction methods affect the validity of the assay in different types of culture media (nutrient broth, commercial milk, normal bovine milk, and heat-sterilized milk). Each tube was inoculated with a full large colony or inoculated from a portion of a single colony (increase in inoculums may increase the chances of detecting previously-existing repeat mutations by PCR, if any). Discrete colonies were used for inoculation to ensure homogeneity of the input profile. They were then passed for three consecutive days, with copy numbers checked from each passage. To see the effect of DNA extraction chemicals on the repeat stability, each sample was divided into two portions, one portion extracted by a specialized kit following recommendations of the manufacturer (Norgen Bioteck Inc, Ontario, Canada) and the other by the phenol/chloroform method.

Repeat stability testing

To test repeat stability, the well characterized reference strains including the two human strains *S. aureus* Lowenstein and Wright, and the noncapsulated variant Smith Compact (CP) were used for serial passaging in bovine milk and nutrient broth (NB), incubated at 37ºC with and without shaker (200rpm). In addition, sporadic bovine mastitis

isolates of *S. aureus* were used for testing variations in the *clfA* R-domain TR copy numbers before and after *in vitro* serial passages in nutrient broth (NB). Nutrient agar (NA) and NB pH 7.2 (Laboratories, Detroit, MI, USA) were used as basic media to prepare cultures before inoculating milk samples.

A. Serial Passage experiment

To examine the effects of different cultural conditions on the *clfA* R-domain repeats, continuous 15x serial passage experiments in commercial milk and in NB were conducted as follows: 5 ml of milk were added to eight disposable sterile 50 ml test tubes (Fisherbrand, Fisher Scientific, Pittsburgh, PA), six of these tubes were inoculated (from a single 24 h old colony of each strain on NA plate) in duplicates with the three strains of *S. aureus*, Smith, Lowenstein, and Wright, two uninoculated tubes were used as negative controls. One set of three inoculated tubes plus an uninoculated control tube were incubated at 37ºC incubator without shaking, and the other set was put at 37ºC shaker condition (200 rpm). Tubes were closed and sealed with parafilm. Two similar sets of NB tubes were also inoculated and incubated under the two culture media conditions for comparison. These cultures (16 tubes) were serially passed for 15x in the milk and NB subculturing after each 24 h period and copy numbers were checked by PCR every two days and confirmed by sequencing. Similarly, single colonies of the 19 bovine mastitis isolates were inoculated in NB and passed for 15 passages, in a similar manner as mentioned above. DNA was immediately isolated from passages P1, P3, P6, P9, P12, and P15. Because homogeneity of the input profile was preferred over quantity, and because it is known that increased mutations are prone to occur in liquid than solid media, single discrete colonies were used for initial inoculation.

B. S. aureus invasion into mammary epithelial cell line (MAC-T cells)

The invasion experiment was carried out as described by Bayles et al., (1998). Prior to each experiment, a single colony each from an overnight NA plate (NA and NB were used to avoid probable effects of rich media on repeats) (Difco laboratories Detroit, MI, USA) of the three *S. aureus* strains, Smith, Lowenstein, and Wright, was inoculated

into 4 ml of NB and grown at 37°C with vigorous shaking for 6 to 8 h. From this 4-ml culture, 100 µl was transferred into 10 ml of NB and incubated overnight (16 h) at 37°C with vigorous shaking. The overnight culture was centrifuged, and the pellet was washed once with sterile phosphate-buffered saline (pH 7.2) and resuspended in 10 ml of invasion medium (see below) to give a cell density of 10^{10} CFU ml^{-1}.

Cell Culture

An established bovine mammary epithelial cell line, designated MAC-T (Huynh et al., 1991) was used for invasion experiments. Composition of the MAC-T cell growth medium was Dulbecco's Modified Eagle Medium (Sigma-Aldrich Ltd, Oakville, Ontario, Canada) 44.5%, RPMI Medium 1640 (1X) with L-glutamine (Invitrogen Canada Inc., Burlington, Ontario, Canada) 44.5%, Fetal Bovine Serum (FBS, Gibco/Invitrogen Inc.) 10%, and 1ml of 100x concentrated antibiotic/antimycotic solution (Invitrogen Inc.), sterilized by filtration using 0.22 µm pore size. Cells were seeded at 6 x 10^4 cells/well and grown for 3 days (or till about 90 % confluence) at 37°C with 5% CO_2. Cells were grown in 12-well tissue culture plates for invasion.

Invasion Assay

The invasion was carried out as described by Bayles et al., (1998). Approximately 16 h prior to invasion experiment the MAC-T cell growth medium was replaced with 1 ml of invasion medium (growth medium without antibiotics or FBS). The morning of the experiment, the medium was removed and MAC-T cells were washed once with the invasion medium and given 1 ml of fresh invasion medium. Appropriate wells of MAC-T cells were then inoculated with 10^7 CFU of washed *S. aureus* and incubated at 37°C with 5% CO_2. After 2 h, supernatants of the cocultures were removed and monolayers were then incubated with lysostaphin (10 µg ml^{-1}; Sigma) in invasion medium for 20 min at 37°C to kill extracellular bacteria (Qazi et al., 2004). After the incubation of cocultures at 37°C with 5% CO_2 for additional 16 h, the supernatants were removed and discarded. MAC-T cell monolayers were washed three times with sterile PBS, treated with 0.25% trypsin in Hanks balanced salt solution (Invitrogen Inc., Burlington, Ontario, Canada) for about 5 mins at 37°C. They were further lysed with 0.025% Triton X-100 in sterile

distilled water. Cell lysates were immediately used for extraction of DNA and used as template for R-domain copy number analysis.

clfA R-domain polymorphism in published sequences of S. aureus

In order to see the extent of polymorphism of *clfA* R-domain among known clinical *S. aureus* strains, and to understand the VNTR profile and structure in different strains, the published sequences of *S. aureus* were checked for R-domain repeats. The Tandem Repeat Finder program (v 4.0) (Benson, 1999) (http://tandem.bu.edu/trf/trf.html) was used to locate and sort the repeats in the R-domain of the *clfA* from genome sequences of all strains listed in Table 2.1. These strains cover a range of *S. aureus* including MSSA, HA-MRSA (hospital acquired MRSA), CA-MSSA (community acquired MSSA), CA-MRSA (community acquired MRSA), and MRSA-VRSA (MRSA-vancomycin resistant *S. aureus*). The nucleotide sequences of the *clfA* from those strains were also aligned using the ClustalW (v 1.8) from the European Bioinformatics Institute web site (http://www.ebi.ac.uk/Tools/clustalw/index.html**),** the DNAstar, and Geneious v 3.5.6 (Biomatter Ltd, Auckland, New Zealand) programs to analyze differences in copy number.

PCR amplified clfA R-domain polymorphism in S. aureus isolates from human infections and bovine mastitis

A collection of 19 sporadic bovine mastitis isolates of *S. aureus* as well as of 24 MSSA isolates from diverse populations of human patients were used for testing variations in the *clfA* R-domain TR copy numbers. Primer pairs flanking the *clfA* R region were designed using DNAstar and Primer3 (Rosen and Skaletsky, 2000) programs on conserved region on the chromosome of *S. aureus* COL after aligning *clfA* sequences obtained from sequenced strains (Table2.1). Sequences were BLAST checked for homologous regions. Primer sequences were designed with the forward primer as 5'CCTGATGAGCCTGGTGAAAT 3' and the reverse primer as 5' TTAGAACCTGACTCGGAATCG 3'. Primers pairs were synthesized by Invitrogen (Invitrogen Inc.). PCR reactions were carried out on TechGene version 12.02A (Techne® Inc. NJ, USA) and the conditions that allowed amplification of single PCR products were

established as 30 cycles each of denaturation at 94ºC for 1 min, annealing at 54ºC for 1 min, extension at 72ºC for 1 min, and a final extension of 72ºC for 5 min. All PCR reactions, including those of human and mastitis isolates, were then run under the same conditions. Fragments were separated on 1% agarose gel by electrophoresis using 1% TBE as running and gel buffer. Differences in R-domain copy numbers were inferred from PCR product sizes on gels. The primers give a product size of 962 bp in *S. aureus* COL that corresponds to 50.5 copies of repeats in the R-domain. The inferred copy numbers of different alleles were confirmed by sequencing PCR products at the Genome Quebec and McGill University Innovation Center.

Determination of alleles was based on differences in the number of the 18bp copies of tandem repeats contained in the R domain of the *clfA* gene. The numerical index of discriminatory power was used to give numeric estimates for the discriminations between strains/isolates. The discriminatory power values (defined as the average probability that the typing system will assign a different type to two unrelated strains randomly sampled in the microbial population of a given taxon) were estimated according to Hunter (1990). The value of (1) indicates the ability to differentiate each isolate, and (0) indicates that all isolates are identical.

Nucleotide sequence analysis and phylogenetic relationships

Chromatograms and text files of the resulting sequences were checked and assembled using Chromas version 2.32 program. Multiple sequence alignments of the strains and isolates, including those used *in vitro* passaging and invasion into MAC-T cell, were done using Chromas and ClustalW from European Biofinformatics Institute sites (http://www.ebi.ac.uk/Tools/clustalw/index.html). Before this analysis, copy numbers of TRs on the sequences were checked with the help of the Tandem Repeat Finder program (Benson, 1999) (http://tandem.bu.edu/trf/trf.html) and confirmed by visual sorting of individual copy numbers. Since the repeats in the R-domain are highly degenerate and the sequences vary at both 3' and 5' ends of the R region, comparative phylogenetic grouping was carried out using whole R-domain sequences including parts of conserved regions with the Neighbor-Joining method built in the Geneious bioinformatics package.

RESULTS

Pilot experiment

The results from the pilot experiment showed that the type of sterilization used for milk samples did not have any effect on repeat numbers (data not shown). Consequently, only commercial milk (Lactantia Purfiltre, Parmalate Canada) was used in later studies. Similarly, the different sizes of bacterial inocula gave same results (data not shown). DNA obtained from different milk samples, inoculum sizes, and DNA extraction methods, all yielded single PCR products of identical sizes 1090 bp which corresponds to 57 copies of the repeat for *S. aureus* Wright. The DNA extracted by the kit and phenol/chloroform methods produced same results. Thus, only the kit was used for later studies.

Repeat stability testing

As shown in Figure 2.1 for Wright, no change in the R-domain repeat copy-number was observed after passages P1 to P15 in milk and NB in 37ºC incubators with or without shaker. Repeats also remained invariant for *S. aureus* Smith and Lowenstein passed under the same conditions (Table 2.2). Similarly, no change in the repeat copy number occurred during serial passage of the 19 mastitis isolates in NB (Table 2.2). Sixteen mastitis isolates had a product size of 1000bp (equivalent to 52 copies); whereas, two other isolates, (760M and 855M) had a product size of 960bp (50 copies), and another isolate (140M) had a 850bp product (44 copies). In addition, there was no change in the repeats after bacterial internalization into MAC-T cells as detected by PCR (Table 2.2). Representative PCR products of the three strains and mastitis isolates in shaker were confirmed by DNA sequencing. The copy numbers calculated from the sequences were identical to the inferred repeat copy number.

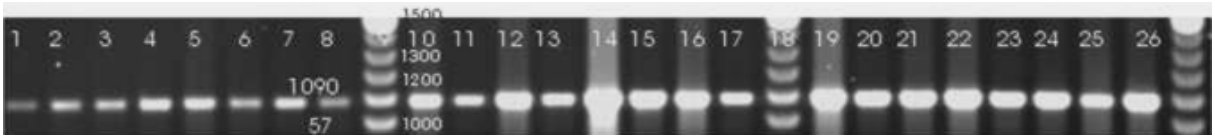


Figure 2.1 PCR product and the corresponding inferred R-domain copy numbers of strain Wright after passage in bovine milk and NB for two weeks at 37ºC incubators with and without shaking. Numbers above and below bands are PCR product sizes and the corresponding inferred copies, respectively. Lanes1=NBP1 incubator, 2 = milkP1 incubator, 3=NBP1 shaker, 4=milk P1 shaker 5=NBP3incubator, 6=milk P3 incubator, 7=NBP3shaker, 8milk P3 shaker, 9=100bp ladder 10=NBP6 incubator, 11=milkP6 incubator, 12=NBP6 shaker, 13=milkP6 shake 14=NBP9 incubator, 15=milkP9 incubator, 16= NBP9 shaker, 17milkP9 shaker, 18= ladder 19=NBP12incubator, 20=milkP12incubator, 21=NBP12shaker, 22=milkP12 shaker 23=NBP15incubator, 24=milkP15incubator, 25=NBP15shaker, 26=milkP15 shaker, 27=ladder (100bp Invitrogen Inc Canada)

Table 2.2. Stability test of R-domain repeats during serial passage experiments and after MAC-T cell invasion.

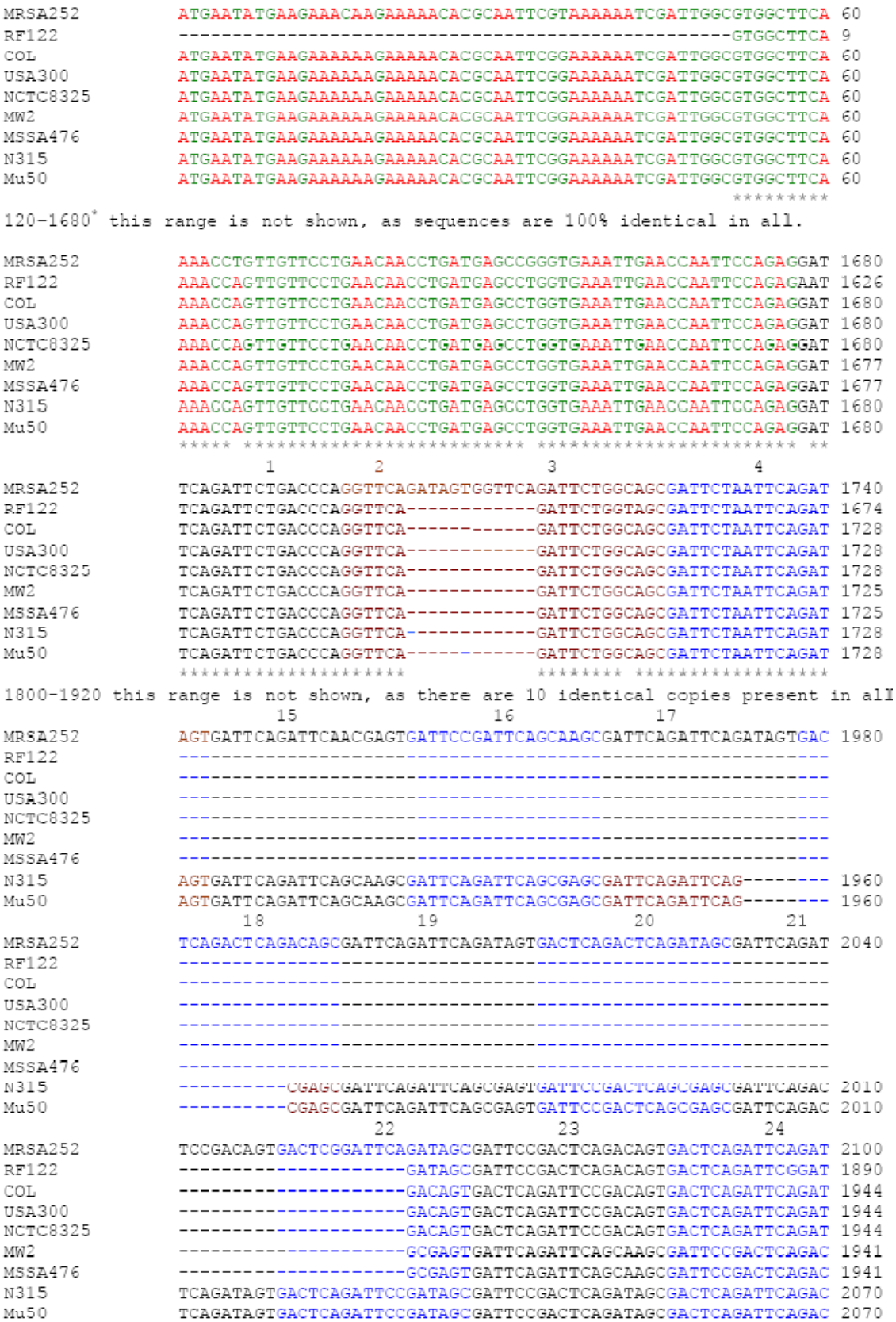
Strain/isolates	**PCR product size**[a](bp) **after passage in**		**copy#**[b]
S. aureus Smith	NB P1, P3, P6, P9, P12, and P15	1250	66
	Milk P1, P3, P6, P9, P12, and P15	1250	66
	Invasion in MAC-T	1250	66
Lowenstein	NB P1, P3, P6, P9, P12, and P15	960	50
	Milk P1, P3, P6, P9, P12, and P15	960	50
	Invasion in MAC-T	960	50
Wright	NB P1, P3, P6, P9, P12, and P15	1090	57
	Milk P1, P3, P6, P9, P12, and P15	1090	57
	Invasion in MAC-T	1090	57
140M[c]	NB P1, P3, P6, P9, P12, and P15	850	44
760M[c] 855M[c]	NB P1, P3, P6, P9, P12, and P15	960	50
Remaining 16 mastitis isolates[d]	NB P1, P3, P6, P9, P12, and P15	1000	52

[a] Only one product size is shown for each amplification, as no changes occurred in repeat copies during passages P1, P3, P6, P9, P12, and P15 of the three strains Smith, Lowenstein, and Wright, in milk and NB incubated with and without shaking, and after MAC-T invasion. [b] The copy

numbers were inferred from the PCR product sizes on the 1% agarose gel and confirmed by sequencing. Since all bands were of identical sizes (Figure 2.1), irrespective of media or conditions, only the shaker results are shown here. [c] These are bovine mastitis isolates; [d] List of other 16 bovine mastitis isolates: by original codes, SA4, SA 92, SA 140, SA 143, SA 144, SA 145, SA 152, SA 209, SA 358, SA 363, SA 428, SA 429, SA 431, SA 439, SA 758, SA 759, SA 760, SA 764, SA 855. For simplicity the locally assigned numbers 1 to 19 were also used.

clfA R-domain polymorphism in published sequences of different strains of S. aureus

In order to see the existence and the degree of polymorphism of the *clfA* R-domain of *S. aureus*, we first determined repeat copy numbers in nine published genomes available at the time (Table 2.1). Eight of these strains were of human origins except for the RF122 mastitis strain. The VNTR analysis revealed a high degree of variation in repeat copy numbers (Figure 2.2 and Tables 2.1 and 2.3), ranging from 46 copies in mastitis strain RF22 to 66 copies in the MRSA252. The virulent strains MSSA476 and USA300, both with identical copy numbers of 49 each, as well as the strain Mu50 with 49.5 copies, all showed almost equal copy-numbers in their R-domains. Similarly, the rest of the strains in this group namely, NCTC8325, COL, and MW2 had similar but not identical copy-numbers of 48, 50.5, and 51 copies, respectively. However, the strain N315 was found to contain 59.5 copies of repeats, and together with the MRSA252 were the only two in these nine strains that possessed higher copy-numbers in their R-domains. The variations in copy-numbers would correlate well with the variations in the lengths of the whole ClfA as can be seen from the nucleotide sequences in Figure 2.2. Thus, eight genotypes were identified in the nine strain sequences studied (Tables 2.1 and 2.3).


MRSA252 ATGAATATGAAGAAACAAGAAAAACACGCAATTCGTAAAAAATCGATTGGCGTGGCTTCA 60
RF122 ----------GTGGCTTCA 9
COL ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
USA300 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
NCTC8325 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
MW2 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
MSSA476 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
N315 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
Mu50 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60
120-1680* this range is not shown, as sequences are 100% identical in all.
MRSA252 AAACCTGTTGTTCCTGAACAACCTGATGAGCCGGGTGAAATTGAACCAATTCCAGAGGAT 1680
RF122 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGAAT 1626
COL AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1680
USA300 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1680
NCTC8325 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1680
MW2 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1677
MSSA476 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1677
N315 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1680
Mu50 AAACCAGTTGTTCCTGAACAACCTGATGAGCCTGGTGAAATTGAACCAATTCCAGAGGAT 1680
1 2 3 4
MRSA252 TCAGATTCTGACCCAGGTTCAGATAGTGGTTCAGATTCTGGCAGCGATTCTAATTCAGAT 1740
RF122 TCAGATTCTGACCCAGGTTCA----------GATTCTGGTAGCGATTCTAATTCAGAT 1674
COL TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1728
USA300 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1728
NCTC8325 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1728
MW2 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1725
MSSA476 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1725
N315 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1728
Mu50 TCAGATTCTGACCCAGGTTCA----------GATTCTGGCAGCGATTCTAATTCAGAT 1728
1800-1920 this range is not shown, as there are 10 identical copies present in all
15 16 17
MRSA252 AGTGATTCAGATTCAACGAGTGATTCCGATTCAGCAAGCGATTCAGATTCAGATAGTGAC 1980
RF122 ----------
COL ----------
USA300 ----------
NCTC8325 ----------
MW2 ----------
MSSA476 ----------
N315 AGTGATTCAGATTCAGCAAGCGATTCAGATTCAGCGAGCGATTCAGATTCAG-------- 1960
Mu50 AGTGATTCAGATTCAGCAAGCGATTCAGATTCAGCGAGCGATTCAGATTCAG-------- 1960
18 19 20 21
MRSA252 TCAGACTCAGACAGCGATTCAGATTCAGATAGTGACTCAGACTCAGATAGCGATTCAGAT 2040
RF122 ----------
COL ----------
USA300 ----------
NCTC8325 ----------
MW2 ----------
MSSA476 ----------
N315 ----------CGAGCGATTCAGATTCAGCGAGTGATTCCGACTCAGCGAGCGATTCAGAC 2010
Mu50 ----------CGAGCGATTCAGATTCAGCGAGTGATTCCGACTCAGCGAGCGATTCAGAC 2010
22 23 24
MRSA252 TCCGACAGTGACTCGGATTCAGATAGCGATTCCGACTCAGACAGTGACTCAGATTCAGAT 2100
RF122 ----------GATAGCGATTCCGACTCAGACAGTGACTCAGATTCGGAT 1890
COL ----------GACAGTGACTCAGATTCCGACAGTGACTCAGATTCAGAT 1944
USA300 ----------GACAGTGACTCAGATTCCGACAGTGACTCAGATTCAGAT 1944
NCTC8325 ----------GACAGTGACTCAGATTCCGACAGTGACTCAGATTCAGAT 1944
MW2 ----------GCGAGTGATTCAGATTCAGCAAGCGATTCCGACTCAGAC 1941
MSSA476 ----------GCGAGTGATTCAGATTCAGCAAGCGATTCCGACTCAGAC 1941
N315 TCAGATAGTGACTCAGATTCCGATAGCGATTCCGACTCAGATAGCGACTCAGATTCAGAC 2070
Mu50 TCAGATAGTGACTCAGATTCCGATAGCGATTCCGACTCAGATAGCGACTCAGATTCAGAC 2070

```
                          25                  26                  27
MRSA252         AGCGACTCAGACTCAGAAAGCGACTCAGATTCAGACAGCGATTCCGACTCAGATAGTGAT 2160
RF122           AGCGATTCCGACTCAGACAGTGACTCAGATTCGGATAGCGATTCCGACTCAGACAGTGAT 1950
COL             AGCGATTCTGACTCAGACAGTGACTCAGATTCAGATAGCGATTCAGATTCAGATAGCGAT 2004
USA300          AGCGATTCTGACTCAGACAGTGACTCAGATTCAGATAGCGATTCAGATTCAGATAGCGAT 2004
NCTC8325        AGCGATTCTGACTCAGACAGTGACTCGGATTCAGATAGCGATTCAGATTCAGATAGCGAT 2004
MW2             AATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGACTCAGATTCCGATAGCGAT 2001
MSSA476         AATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGACTCAGATTCCGATAGCGAT 2001
N315            AGCGATTCTGACTCAGACAGCGATTCTGACTCAGACAGTGACTCAGATTCCGATAGCGAT 2130
Mu50            AGCGATTCTGACTCAGACAGCGATTCTGACTCAGACAGTGACTCAGATTCCGATAGCGAT 2130
                *  ** ** ** ***** ** ** ** ** ** ** ** ** ** ** ** ** ** ***
                     28          29                  30          31
MRSA252         TCAGATTCAGACAGCGACTCAGACA------GTGTTTCAGATTCAGACAGTGACTCAGAC 2214
RF122           TCAGATTCGGATAGCGATTCCGAC---------------------------------- 1974
COL             TC------------------------------------------------------- 2006
USA300          TC------------------------------------------------------- 2006
NCTC8325        TC------------------------------------------------------- 2006
MW2             TCTGACTCAGACAGTGACTCAGAT-----------------TCAGATAGCGATTCAGAT 2043
MSSA476         TC------------------------------------------------------- 2003
N315            TCCGACTCAGACAGTGACTCAGATTCCGATAGCGATTCCGACTCAGACAGTGACTCAGAT 2190
Mu50            TC------------------------------------------------------- 2132
                **
                               32                  33                  34
MRSA252         TCGGATAGTGAATCAGACTCAGACAGCGACTCAGATTCAGACAGCGACTCAGACTCGGAT 2274
RF122           ----------------------------------TCAGACAGTGACTCAGATTCGGAT 1998
COL             ----------AGATTCCGACAGTGATTCCGACTCAGACAGCGATTCTGACTCCGAC 2052
USA300          ----------AGATTCCGACAGTGATTCCGACTCAGACAGCGATTCTGACTCCGAC 2052
NCTC8325        ----------AGATTCCGACAGTGATTCCGACTCAGACAGCGATTCTGACTCCGAC 2052
MW2             TCAGATAGCGATTCAGATTCCGACAGTGATTCCGACTCAGACAGCGATTCTGACTCCGAC 2103
MSSA476         ----------AGATTCCGACAGTGATTCCGACTCAGACAGCGATTCTGACTCCGAC 2049
N315            TCCGATAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGATTCCGACTCAGAT 2250
Mu50            ------------------------------------------------------------
                          35                  36                  37
MRSA252         AGTGAATCAGACTCAGACAGTGACTCAGACTCAGACAGCGACTCAGACTCAGATAGCGAC 2334
RF122           AGCGATTCCGACTCAGACAGTGACTCAGATTCGGATAGCGATTCAGATTCAGCAAGCGAT 2058
COL             AGTGATTCCGACTCAGACAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGAT 2112
USA300          AGTGATTCCGACTCAGACAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGAT 2112
NCTC8325        AGTGATTCCGACTCAGACAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGAT 2112
MW2             AGTGATTCCGACTCAGACAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGAT 2163
MSSA476         AGTGATTCCGACTCAGACAGCGATTCAGATTCCGACAGTGATTCCGACTCAGATAGCGAT 2109
N315            AGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCTGACTCAGACAGTGAC 2310
Mu50            ---------------------------------------------TGACTCAGACAGTGAC 2148
                                                              ** ****  ** **
                   38                  39                  40
MRSA252         TCAGATTCAGATAGCGATTCAGACTCAGACAGCGACTCAGACTCAGATAGCGATTCCGAC 2394
RF122           TCAGATTCCGATAGTGATTCAGATTCAGCAAGCGATTCAGATTCCGACAGTGATTCCGAC 2118
COL             TCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT 2172
USA300          TCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT 2172
NCTC8325        TCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT 2172
MW2             TCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT 2223
MSSA476         TCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT 2169
N315            TCAGATTCCGATAGCGATTCAGATTCAGACAGTGATTCAGACTCAGATAGCGATTCAGAT 2370
Mu50            TCAGATTCCGATAGCGATTCAGATTCAGACAGTGATTCAGACTCAGATAGCGATTCAGAT 2208
                ** ** ** ***** ** ***** ****  ** ** ***** ** ** ** ***** **
                  41          42                  43                  44
MRSA252         TCGGATAGCGACTCAGACTCAGATAGCGATTCCGACTCGGATAGCGATTCAGAATCAGAT 2454
RF122           TCGGACAGCGACTCAGACTCAGACAGTGA---------------------------- 2147
COL             TCAGATAGCGATTCAGATTCCGACAGTGA---------------------------- 2201
USA300          TCAGATAGCGATTCAGATTCCGACAGTGA---------------------------- 2201
NCTC8325        TCAGATAGCGATTCAGATTCCGACAGTGA---------------------------- 2201
MW2             TCAGATAGCGATTCAGATTCCAACAGTGA---------------------------- 2252
MSSA476         TCAGATAGCGATTCAGATTCCAACAGTGA---------------------------- 2198
N315            TCCGACAGTGACTCAGACTCAGACAGCGA---------------------------- 2399
Mu50            TCCGACAGTGACTCAGACTCAGACAGCGA---------------------------- 2237
                ** ** ** ** ***** **  * ** **
```

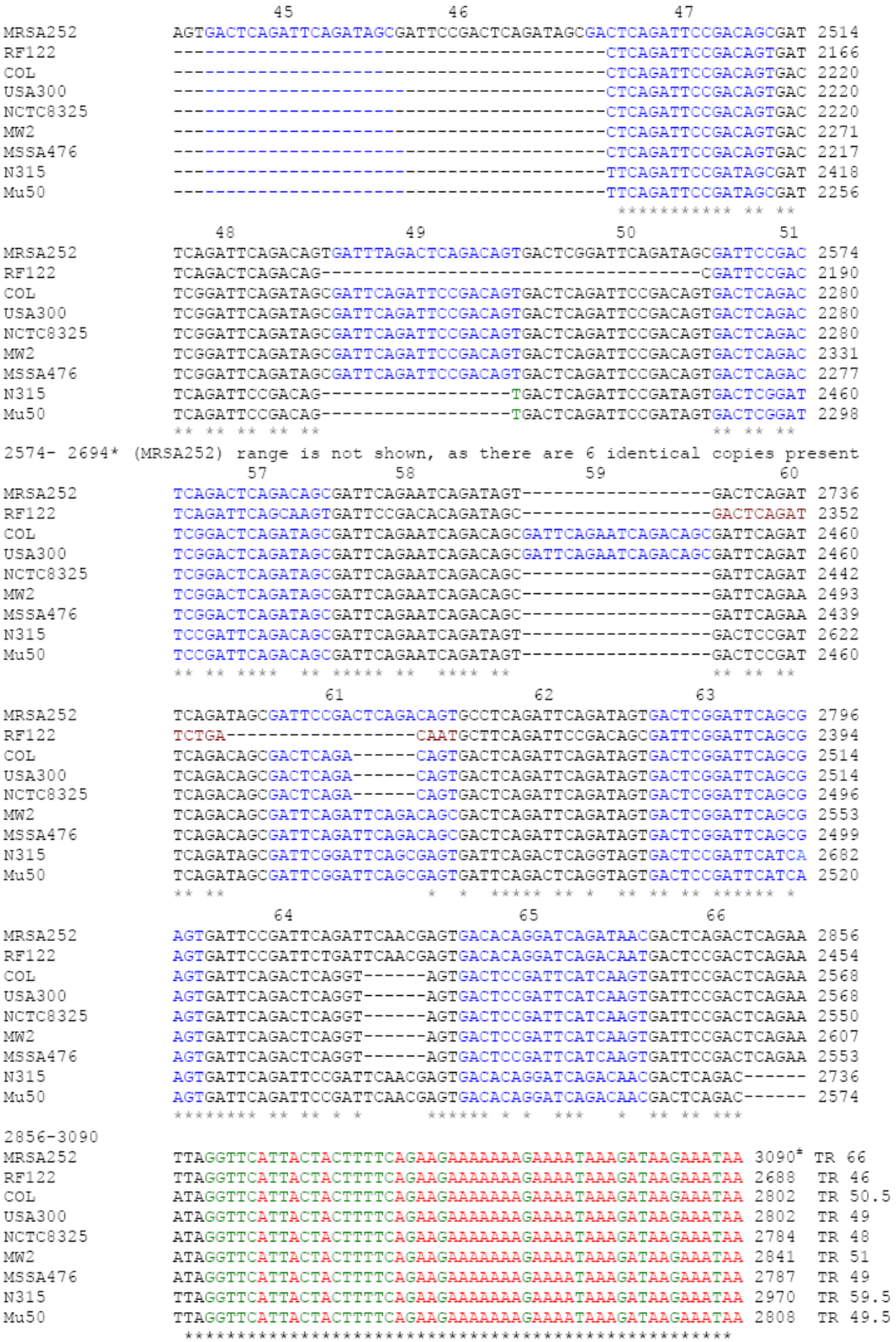

45 46 47
MRSA252 AGTGACTCAGATTCAGATAGCGATTCCGACTCAGATAGCGACTCAGATTCCGACAGCGAT 2514
RF122 CTCAGATTCCGACAGTGAT 2166
COL CTCAGATTCCGACAGTGAC 2220
USA300 CTCAGATTCCGACAGTGAC 2220
NCTC8325 CTCAGATTCCGACAGTGAC 2220
MW2 CTCAGATTCCGACAGTGAC 2271
MSSA476 CTCAGATTCCGACAGTGAC 2217
N315 TTCAGATTCCGATAGCGAT 2418
Mu50 TTCAGATTCCGATAGCGAT 2256
48 49 50 51
MRSA252 TCAGATTCAGACAGTGATTTAGACTCAGACAGTGACTCGGATTCAGATAGCGATTCCGAC 2574
RF122 TCAGACTCAGACAG CGATTCCGAC 2190
COL TCGGATTCAGATAGCGATTCAGATTCCGACAGTGACTCAGATTCCGACAGTGACTCAGAC 2280
USA300 TCGGATTCAGATAGCGATTCAGATTCCGACAGTGACTCAGATTCCGACAGTGACTCAGAC 2280
NCTC8325 TCGGATTCAGATAGCGATTCAGATTCCGACAGTGACTCAGATTCCGACAGTGACTCAGAC 2280
MW2 TCGGATTCAGATAGCGATTCAGATTCCGACAGTGACTCAGATTCCGACAGTGACTCAGAC 2331
MSSA476 TCGGATTCAGATAGCGATTCAGATTCCGACAGTGACTCAGATTCCGACAGTGACTCAGAC 2277
N315 TCAGATTCCGACAG TGACTCAGATTCCGATAGTGACTCGGAT 2460
Mu50 TCAGATTCCGACAG TGACTCAGATTCCGATAGTGACTCGGAT 2298
2574- 2694* (MRSA252) range is not shown, as there are 6 identical copies present
57 58 59 60
MRSA252 TCAGACTCAGACAGCGATTCAGAATCAGATAGT GACTCAGAT 2736
RF122 TCAGATTCAGCAAGTGATTCCGACACAGATAGC GACTCAGAT 2352
COL TCGGACTCAGATAGCGATTCAGAATCAGACAGCGATTCAGAATCAGACAGCGATTCAGAT 2460
USA300 TCGGACTCAGATAGCGATTCAGAATCAGACAGCGATTCAGAATCAGACAGCGATTCAGAT 2460
NCTC8325 TCGGACTCAGATAGCGATTCAGAATCAGACAGC GATTCAGAT 2442
MW2 TCGGACTCAGATAGCGATTCAGAATCAGACAGC GATTCAGAA 2493
MSSA476 TCGGACTCAGATAGCGATTCAGAATCAGACAGC GATTCAGAA 2439
N315 TCCGATTCAGACAGCGATTCAGAATCAGATAGT GACTCCGAT 2622
Mu50 TCCGATTCAGACAGCGATTCAGAATCAGATAGT GACTCCGAT 2460
61 62 63
MRSA252 TCAGATAGCGATTCCGACTCAGACAGTGCCTCAGATTCAGATAGTGACTCGGATTCAGCG 2796
RF122 TCTGA CAATGCTTCAGATTCCGACAGCGATTCGGATTCAGCG 2394
COL TCAGACAGCGACTCAGA CAGTGACTCAGATTCAGATAGTGACTCGGATTCAGCG 2514
USA300 TCAGACAGCGACTCAGA CAGTGACTCAGATTCAGATAGTGACTCGGATTCAGCG 2514
NCTC8325 TCAGACAGCGACTCAGA CAGTGACTCAGATTCAGATAGTGACTCGGATTCAGCG 2496
MW2 TCAGACAGCGATTCAGATTCAGACAGCGACTCAGATTCAGATAGTGACTCGGATTCAGCG 2553
MSSA476 TCAGACAGCGATTCAGATTCAGACAGCGACTCAGATTCAGATAGTGACTCGGATTCAGCG 2499
N315 TCAGATAGCGATTCGGATTCAGCGAGTGATTCAGACTCAGGTAGTGACTCCGATTCATCA 2682
Mu50 TCAGATAGCGATTCGGATTCAGCGAGTGATTCAGACTCAGGTAGTGACTCCGATTCATCA 2520
64 65 66
MRSA252 AGTGATTCCGATTCAGATTCAACGAGTGACACAGGATCAGATAACGACTCAGACTCAGAA 2856
RF122 AGTGATTCCGATTCTGATTCAACGAGTGACACAGGATCAGACAATGACTCCGACTCAGAA 2454
COL AGTGATTCAGACTCAGGT AGTGACTCCGATTCATCAAGTGATTCCGACTCAGAA 2568
USA300 AGTGATTCAGACTCAGGT AGTGACTCCGATTCATCAAGTGATTCCGACTCAGAA 2568
NCTC8325 AGTGATTCAGACTCAGGT AGTGACTCCGATTCATCAAGTGATTCCGACTCAGAA 2550
MW2 AGTGATTCAGACTCAGGT AGTGACTCCGATTCATCAAGTGATTCCGACTCAGAA 2607
MSSA476 AGTGATTCAGACTCAGGT AGTGACTCCGATTCATCAAGTGATTCCGACTCAGAA 2553
N315 AGTGATTCAGATTCCGATTCAACGAGTGACACAGGATCAGACAACGACTCAGAC 2736
Mu50 AGTGATTCAGATTCCGATTCAACGAGTGACACAGGATCAGACAACGACTCAGAC 2574
2856-3090
MRSA252 TTAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 3090# TR 66
RF122 TTAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2688 TR 46
COL ATAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2802 TR 50.5
USA300 ATAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2802 TR 49
NCTC8325 ATAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2784 TR 48
MW2 ATAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2841 TR 51
MSSA476 ATAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2787 TR 49
N315 TTAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2970 TR 59.5
Mu50 TTAGGTTCATTACTACTTTTCAGAAGAAAAAAAGAAAATAAAGATAAGAAATAA 2808 TR 49.5

Figure 2.2. Multiple sequence alignment of *clfA* R-domains from published *S. aureus* obtained using the ClustalW program, v.1.82, http://www.ebi.ac.uk/Tools/clustalw/index.html. Repeat regions are identified by numbers above each unit. *Regions with conserved sequences and those with identical repeats in all strains are not shown. ±Variation in repeats corresponds to variations in the protein length.

Amplified clfA R-domain polymorphism in S. aureus of human and animal origins

To further determine the existence of variation in the *clfA* R-domain of *S. aureus*, 24 isolates from human infections (all MSSA, in contrast to MRSA in published 8 human strains) and 19 mastitis isolates were also screened. Similar to the *clfA* R-domain variations revealed in published sequences above, there was also a very high degree of polymorphism among the amplified R-domain lengths between *S. aureus* isolates and strains from human infections. As shown in Figure 2.3A, the three strains Smith, Lowenstein, and Wright had significant differences in lengths of the R-domain; total PCR product sizes were 1250, 960, and 1090bp which corresponded to 66, 50, and 57 copies of the repeats, respectively. However, analysis of the MSSA isolates from blood and SF of patients revealed much higher levels of repeat polymorphism ranging from 36 to 73 copies (Figure 2.3B). These were divided into 15 genotypes A to O (Table 2.3). Except for five, all others were represented by single isolates. Genotype I was associated with the highest numbers of isolates (n=5, plus Wright) and had 57 copies of repeats identical to that of strain Wright. Six genotypes (J, K, L, M, N, and O) represented by single isolates, except for N, contained over 60 copies of repeats in their R-domains. Isolates of MSSA genotypes A to I (with the exception of type B which had 36 copy), all had copy numbers between 44-57. Furthermore, analysis of the banding patterns of the sporadic bovine mastitis isolates in Figure 2.3C showed that they were divided into three genotypes; the vast majority of isolates i.e., 16 out of the 19 isolates formed a single genotype (X) with 52 copies of the repeat while only one, and two isolates (with copy numbers 44 and 50) belonged to types Y and Z, respectively (Table 2.3). Therefore, no significant genetic heterogeneity at the R-domain was observed in *S. aureus* isolated from bovine mastitis, indicating the clonal nature of mastitis isolates tested.

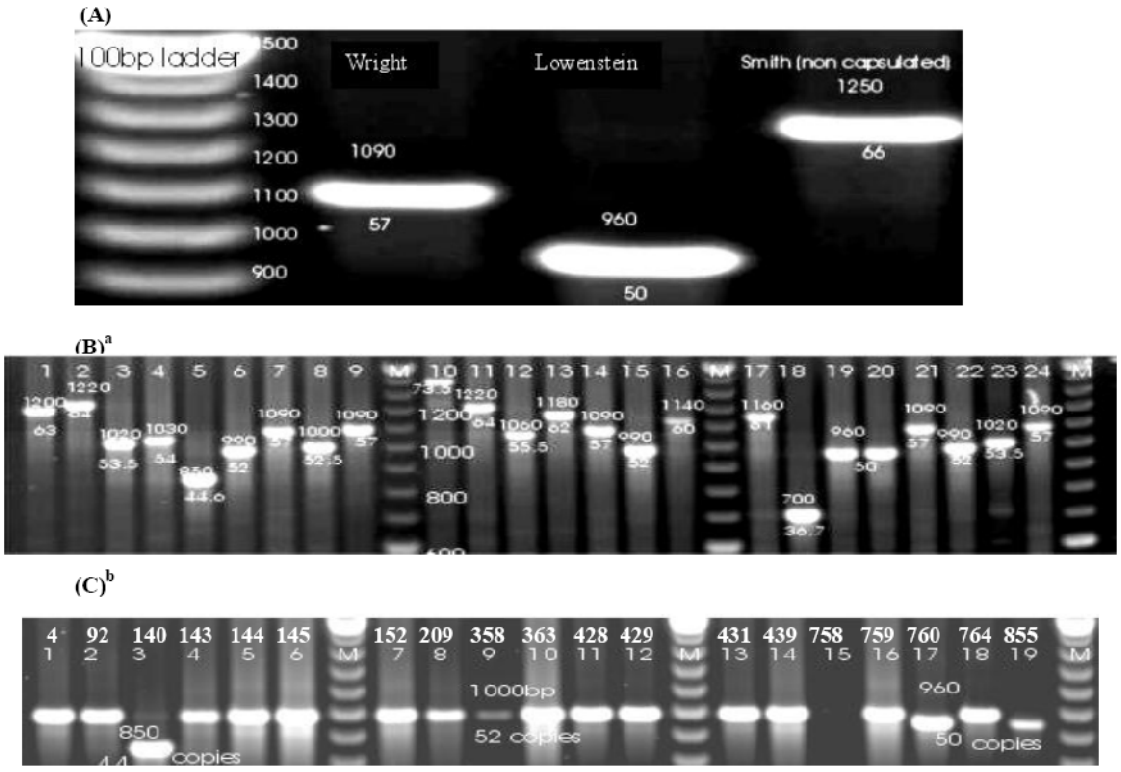


Figure. 2.3. A, B. and C. PCR Amplified *clfA* R-domains in *S. aureus* isolates from (A, B) human infections and (C) bovine mastitis. In all gels, PCR products sizes were indicated above bands, and the corresponding copy numbers are shown below the bands. All PCR and gel conditions were identical. M= 100bpp ladder, all PCR products were measure by 100bp ladder added at 0.1ug per mm lane width, Invitrogen Canada, Inc. [a] Lane numbers 1-24 represent isolates from human infections (1H to 24H).[b] The original code numbers of mastitis isolates shown in bold were used in text associated with letter M for mastitis; numbers from 1 to 19 were locally assigned. [a] Sequence 14H was repeated to confirm accuracy of clustering.

Table 2.3. *clfA* genotypes and the Discriminatory power values for *S. aureus* isolates/strains of human and bovine mastitis origins. The two-sided arrows indicate the low- and high-copy ranges within human *S. aureus,* respectively. Alleles were assigned based on copy number differences; single copy difference was termed as independent allele. Genotypes are given letters A to W.

Human MSSAs			MRSA and Reference strains			Mastitis		
Isolates (24)	**Copy**	**Geno types**[a] (15)	**Isolates** (9 plus 3)	**Copy**	**Geno types** (8)	**Isolates** (19)	**Copy**	**Geno types** (3)
H18	36	A	Mu50	49.5	P	4	52	X16
H5	44	B	RF122	46	Q	92	52	X16
H19, H20	50	C3	COL	50.5	R	140	44	Y
H6, H15, H22	52	D3	USA300	49	S2	143	52	X16
H8	52.5	E	NCTC8325	48	T	144	52	X16
H3, H23	53.5	F2	MW2	51	U	145	52	X16
H4	54	G	MSSA476	49	S2	152	52	X16
H12	55.5	H	Lowenstein	50	C3	209	52	X16
H7,H9,H14, H21, H24	57	I6				358	52	X16
H16	60	J	Smith	66	V2	363	52	X16
H17	61	K				428	52	X16
H13	62	L	Wright	57	I6	429	52	X16
H1	63	M	MRSA252	66	V2	431	52	X16
H2, H11	64	N2	N315	59.5	W	439	52	X16
H10	73.5	O				758	52	X16
						759	52	X16
						760	50	Z2
						764	52	X16
						855	50	Z2

[a] The numbers associated to genotypes indicate number of isolates of that type.

The 55 R-domain sequences, both from published sequences of different strains of *S. aureus* and amplified *clfA* R-domains in *S. aureus* isolates of human and animal origins, formed two main

groups of related strains based on the repeat copy-numbers and sequence data of their R-domains. A low-copy-number group comprised of the 80% (44) of the R-domains studied that included all mastitis and the high-virulence, the invasive, and the CA-MRSA with related copy-numbers in the narrow range of 44 to 57 copies, in addition to the genotype B from human (isolate H5) with the lowest copy number of 36. A high-copy-number group consisted of only ten members of isolates, mainly from the MSSA isolates, and had repeat-copies in the range of 59.5 to 73.

Discriminatory power of clfA R-domain

The discriminatory power was estimated for *S. aureus* isolates/strains from human infections or from mastitis (Table 2.3). The discriminatory power was higher for human isolates/strains both in terms of genotypes (23) as well as the index value of discriminatory power (ID of 0.97). Further, the nine sporadic sequenced strain collections were could be divided into eight different genotypes with an ID of 0.97. The marker could identify three genotypes (one dominant type and two minor types) in the isolates from mastitis collection tested with an ID of 0.30. In overall, the resolution for all 55 *clfA* R-domains was 27 genotypes with an ID of 0.90.

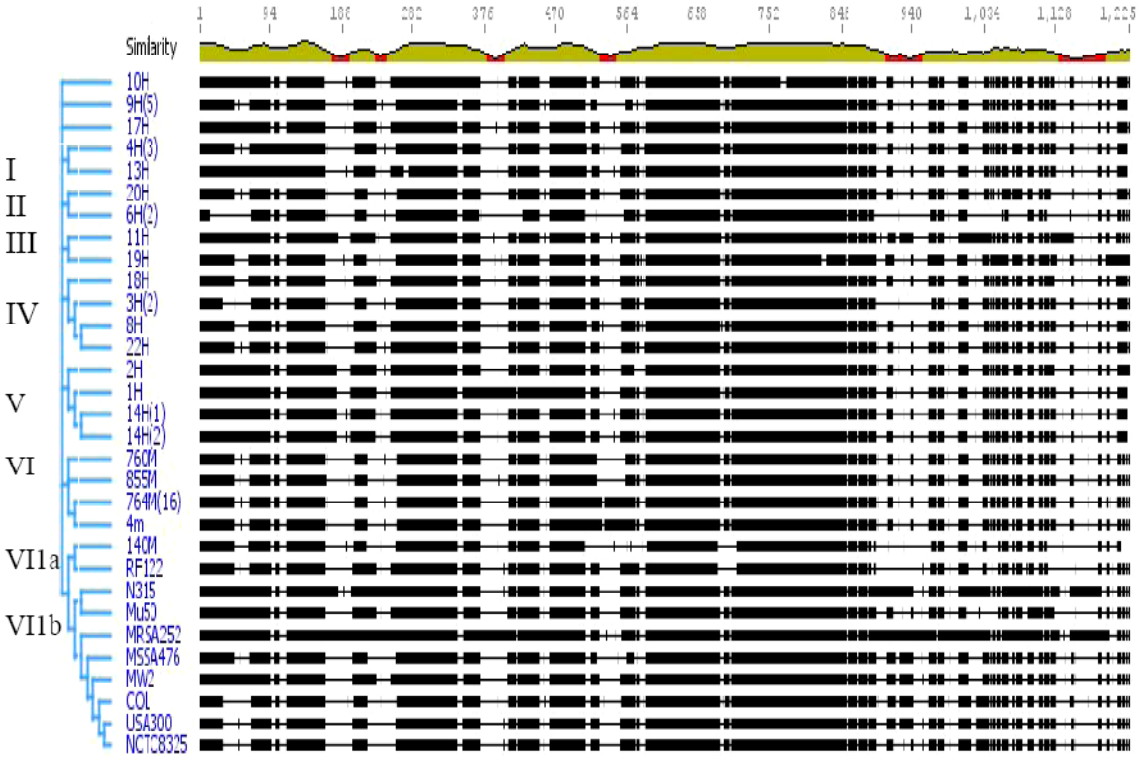
Sequence alignment patterns and phylogenetic relationships

The repeats in the R-domain of the *clfA* are highly degenerate. They vary considerably in the beginning and towards the ends of the repeat regions even between very closely related strains, and perfect only in the middle. Thus, a multiple nucleotide sequence alignments were made to identify the overall sequence patterns among strains. The alignment pattern combined with a basic guide tree is presented in Figure 2.4A. Related groups showed significant similarities in the overall nucleotides. These were distinguishable by a guide tree and into eight distinct groupings labeled as I, II, III, IV, V, VI, VIIa, and VIIb. A representative isolate (9H5) of the major MSSA genotype I that contained five isolates as well as the strain Wright, had formed an independent lineage as did the other two isolates 10H, and 17H. Isolates and strains of groups I, II, III, IV, and V, were highly similar in sequence, particularly, within than between groups. Similarly, all MRSA isolates (VIIb) showed high similarity patterns among related groups. For instance, the HA-associated strains, N315, Mu50 and MRSA252 were similar in sequence, as were strains belonging to the community-associated group, MSSA476,

MW2, USA300 as well as the strains COL and NCTC8325. Overall, except for the MRSA252 which was unique within group VIIb subcluter, all MRSAs had similar sequences. Unlike isolates from human infections, those of the three mastitis genotypes showed much higher nucleotide sequence homology and were separable into two distinct but closely related groups. The dominant group (VI) comprised of two identical alignment pattern groups; in one, the majority of mastitis isolates (16) represented by 764M showed sequence homology to isolate 4M, in the other, isolates 760M and 855M were shown as identical twins. In the minor group (VIIa) isolate 140M was more closely related to the sequenced RF122 strain than to other isolates.

Phylogenetic analysis based on whole nucleotide sequences of the R-domain with parts of the 3' and 5' conserved regions was carried out using Neighbor-Joining method (Tamura-Nei model). As shown in Figure 2.4B, the overall distribution of strains was congruent; all groups shown in Figure 2.4A, were also identified. The nodes had a significantly high percentage of consensus support. Consistent with the sequence alignment patterns in Figure 2.4A, the CA- and HA-associated strains were grouped together more closely related to each other than MSSAs, and were placed in two main subclusters within group VIIb. The MSSA were grouped in five different clusters and independent three lineages (9H, 10H, and 17H). Similarly, mastitis isolates were clustered in two separate groups; VI, which contained almost all of the isolates (Genotypes X and Z), and VIIa in which the isolate 140M (Genotype Y) was grouped with the sequenced RF122 strain closely related to MRSA group.

(A)


1
94
188
282
376
470
564
658
752
846
940
1,034
1,128
1,225
Similarity
10H
9H(5)
17H
4H(3)
13H
20H
6H(2)
11H
19H
18H
3H(2)
8H
22H
2H
1H
14H(1)
14H(2)
760M
855M
764M(16)
4m
140M
RF122
N315
Mu50
MRSA252
MSSA476
MW2
COL
USA300
NCTC8325
I
II
III
IV
V
VI
VI1a
VI1b

(B)

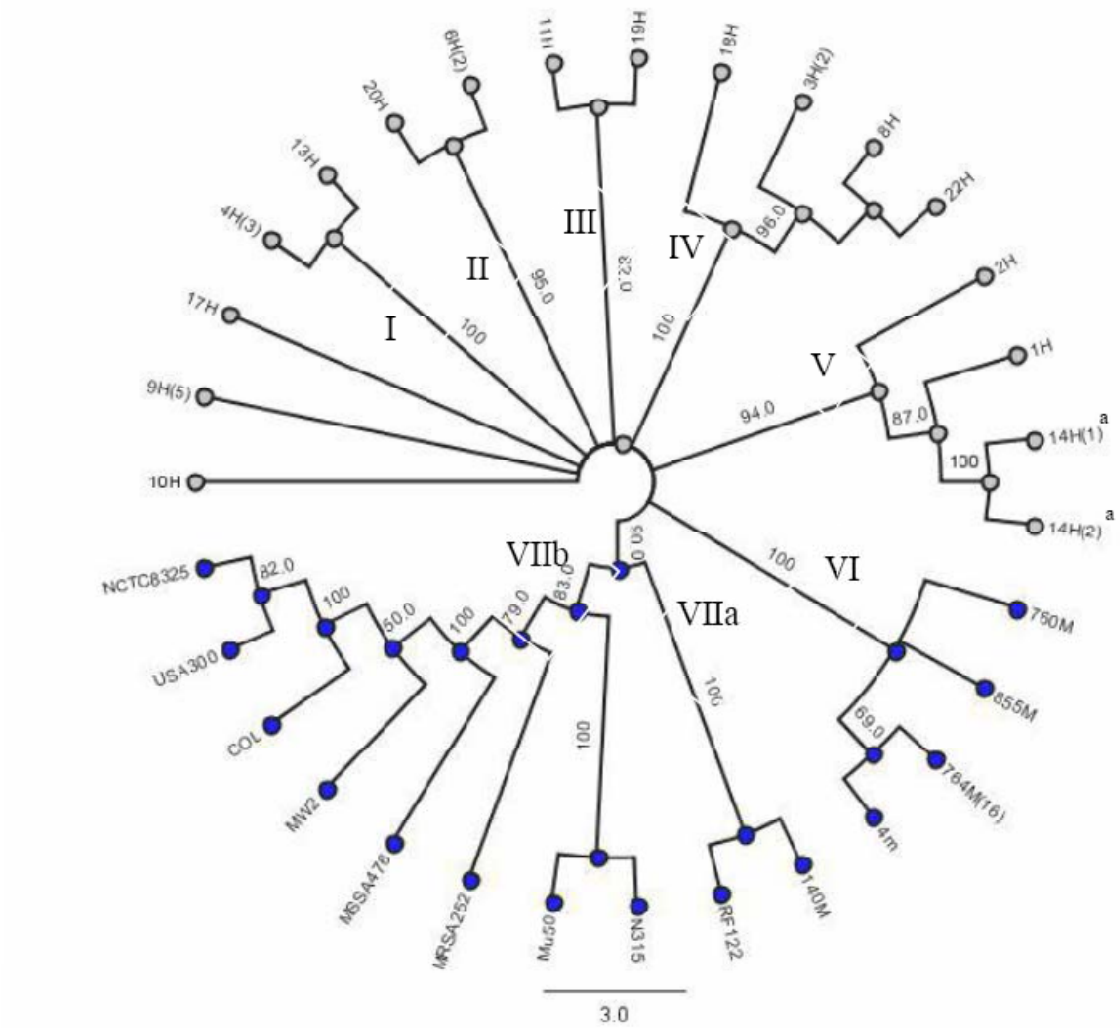


Figure 2.4**A:** Global sequence alignments with a combined guide-tree based on Neighbor-Joining using R-domain nucleotide sequences of representative genotypes from human (n=15), mastitis (n= 5), and the published sequences of *S. aureus* (n=9). Sequence similarity is indicated by the green color. For published sequences standard abbreviations were used. For isolates letters associated by numbers indicate isolate origin (H for human and M for mastitis) and numbers are isolate codes. **B.** Phylogenetic tree based on Neighbor-Joining program in Geneious bioinformatics package (v. 3.5.6) (Distances were estimated by the Tamura-Nei model, bootstrapping100, threshold of 50%) using the R-domain sequences. Blue nodes in cluster VI, VIIa, and VIIb indicate distinct groupings of mastitis and MRSAs strains.

DISCUSSION

This study has shown that *clfA* repeats were stable under different growth media and conditions that are reflective of the host. We have shown that most (80%) *S. aureus* from human infections and mastitis had formed a low-copy group with respect to the number of repeats in their *clfA*-R-domains. Groupings based on copy-numbers and on nucleotide sequences of the R-domain were in agreement and could differentiate isolates of human origin and subclonal populations of mastitis. We have also shown that human isolates were highly polymorphic, while those of mastitis were homogenous. Finally, existence of variant repeats in strains from human and the dominance of a clonal motif in mastitis may imply that a specific selection has occurred in the mammary glands.

VNTR systems have been applied for typing medically important bacteria even without supportive sequencing (Reviewed in van Belkum, 2007). In recent years, these systems employed repeats in intergenic regions. Nevertheless, these regions could be affected due to many factors such as mutations involving frame-shift, recombination, and/or indels that are known to occur in repeated regions. Alternatively, intragenic repeats that have biological significances, especially from relatively stable genes, could be explored for typing. The importance of using coding repeats has received less attention so far, except for a few such as *spa* typing (Koreen et al., 2004). However, the importance of *spa* repeat variation has not been determined, and so far, there is no relationship established between the number of octapeptide repeats and the host species from where strains were isolated (Atkins et al., 2008). Thus, analysis of coding repeat domains useful for both typing and addressing host-specialization are needed.

The *clfA* has several properties of a reasonable marker for typing. It codes for the clumping factor A, a universal core adhesin that mediates virulence in several deep-seated infection models including rat experimental endocarditis (Moreillon et al., 1995), murine sepsis and septic arthritis (Josefsson et al., 2001), and rabbit infective endocarditis (Vernachio et al., 2003). Furthermore, we have shown that this region is stable under different conditions tested. The two incubations at 37ºC, with and without shaker, were set up to test the effect of oxygenation on repeat instability, since oxygen tension has been reported to regulate generation of reactive oxygen species that cause oxidative

modification of DNA leading to transformation from the bacillary to the coccoid form of. *H. pylori* (Park et al., 2004), repeat instability in *E. coli* K12 *tonB* gene (Yamamura et al., 2003), and substitutive mutations in *Streptococcus pneumoniae* (Pericone et al., 2002). Stability was also confirmed by culturing bacteria in thioglycolate media and in carbon dioxide incubators (data not shown). Similarly, since milk components, or metabolic products like acids could have significant effects on gene regulation (Chang et al., 2006), three types of milk samples were tested in a pilot experiment, and had no effect on repeats. More importantly, serial passage of the bacteria in milk and their invasion into the mammary epithelial cells did not affect the repeat copies.

To demonstrate the usefulness of the *clfA* R-domain in typing and grouping of *S. aureus*, published *clfA* sequences from completed *S. aureus* genomes available at the time as well as isolates from human infections and bovine mastitis were analyzed for variations in copy number and nucleotide sequence. MRSA strains, particularly the CA-MRSA group analyzed had shown considerable similarities in copy number and sequence patterns, except for the MRSA252 (Figures 2.2 and 2.4A). Similarly, strains from different hosts were grouped into respective clusters (Figure 2.4B) that were in general agreement with others' findings. Consistent with our results, a recent phylogenetic analysis revealed the close relationships among staphylococcal strains COL, NCTC8325 and USA300 and the relative evolutionary distance to strains MRSA252, MSSA476, MW2, Mu50, N315, and RF122 within the cluster (Baba et al., 2008). It has been reported that MRSA252 was the most genetically diverse strain sequenced (Holden et al., 2004). This would explain why it was relatively distinct from other two HA-MRSA strains, Mu50 and N315 (Figure 2.4). Similarly, grouping of isolates of the three mastitis genotypes into only two subclusters related to human associated strains is in agreement with findings of van Leeuwen et al., (2005) that mastitis strains co-segregated with bacteremia-associated strains from humans. The mastitis isolate (140M) had very similar alignment patterns and copy number to the known mastitis strain RF122 indicating their identity, while all other mastitis isolates formed an independent but related subclonal group as identified by both sequence and copy-number. The human MSSA isolates formed four closely linked groups flanked by clusters of the mastitis and hospital strains.

These data indicated the high-level sensitivity and discriminatory power of the R-domain in differentiating isolates.

The marker's discriminatory power, as determined by the number of possible genotypes as well as by the indexed value for discrimination, was higher (Table 2.3) particularly among different groups of human *S. aureus*. It could resolve 23 genotypes only among human strains/isolates (35) analyzed with high ID (0.97). This is slightly higher than the results of whole region sequence obtained by Kuhn et al., (2007) in related and diverse isolates. Furthermore, resolution of genotypes and MRSA groupings were better than that obtained by *clfB* sequencing (17 types) of the 36 (11 MRSA and 25 MSSA) diverse strains representing the breadth of genetic variation in *S. aureus* (Koreen et al., 2005); however, both markers had a comparable IDs. Eight genotypes were identified in the nine MRSA strains analyzed in this study (ID of 0.97). The marker could identify three genotypes in the mastitis isolates tested (ID of 0.30). Increasing the number of mastitis isolates might have increased the number of genotypes or identified more clonal types; however, compared to the polymorphic nature of human associated *S. aureus*, the clonality of mastitis isolates was evident in the isolates tested. In overall, the resolution in all 55 *clfA* R-domains was 27 genotypes with an ID of 0.90. This result was similar, in terms of the number of genotypes, to the best performing methods PFGE (28 genotypes and ID 96.4) and spa typing (29 types and ID of 97.3) (Koreen et al., 2004); however, the number and diversity of our strains were higher.

It is also interesting to find that the strains and isolates studied formed two groups with respect to the length of their R-domains that reflected the length of the whole ClfA proteins. Both groups showed variations within a defined range of 13 copies. The ten members of the high-copy-number group, all MSSA isolates except three, had repeat-copies in the range of 59.5 to 73, whereas the vast majority the tested strains/isolates belonged to the low-copy-number group with 44 to 57 copies. The latter group contained (groups are indicated in Table 2.3 by arrows) mastitis and MRSA including the hypervirulent, invasive, MRSA-VRSA, and the CA-MRSAs strains studied. The significant larger number of isolates/strains (44 out of the total of 55) belonging to the lower-copy-number group is interesting and it might be related to the invasiveness of

these bacteria. The exact relationship between the copy number and the invasiveness of *S. aureus* is worthy of further investigation.

Our findings from this study may contribute to a better understanding of *clfA* repeat copy-number ranges in differentiating virulent clones and in identifying basis for selection in different organs. This would add more insights on the colonization and adherence properties to different host-tissues; and thus on the mechanisms of superior epidemicity and dominance of specific clones. *clfA-R* typing will have a potential as a complementary typing method for tracking local and global strain profiles and examining *S. aureus* transmission, ecology, and evolution. The simplicity, sensitivity, and reproducibility of the method described in this study make it a valuable tool which can be used adequately by almost any typing laboratory.

ACKNOWLEDGEMENTS

This study was supported by funding from a discovery grant from Natural Science and Engineering Research Council of Canada and from the Canadian Mastitis Research Network.

REFERENCES

Atkins KL, Burman JD, Chamberlain ES , Cooper JE, Poutrel B, Bagbya S, Jenkins AT, Feil EJ, van den Elsen JM. 2008. *S. aureus* IgG-binding proteins SpA and Sbi: Host specificity and mechanisms of immune complex formation**.** Mol Immunol 45**:**1600–1611.

Ayraud S, Janvier B, Salaun L, Fauchere JL. 2003. Modification in the *ppk* Gene of *Helicobacter pylori* during single and multiple experimental murine infections. Infect Immun 71: 1733–1739.

Baba T, Bae T, Schneewind O, Takeuchi F, Hiramatsu K. 2008. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of

Staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. J Bacteriol 190:300-310.

Barrio M, Rainard P, Poutrel B. 2003. Milk complement and the opsonophagocytosis and killing of *Staphylococcus aureus* mastitis isolates by bovine neutrophils. Microbial Path *34: 1–9*

.

Benson G.1999.Tandem Repeats Finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573-580.

Bayliss CD, Field D, Moxon ER. 2001. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis.* J Clin Invest 107:657-662.

Bayles KW, Wesson CA, Liou LE, Fox LK, Bohach GA, Trumble WR. 1998. Intracellular *Staphylococcus aureus* escapes the endosome and induces apoptosis in epithelial cells. Infect Immun 66:336-342.

Chang W, Toghrol F, Bentley WE. 2006. Toxicogenomic response of *Staphylococcus aureus* to peracetic acid. Environ Sci Technol 40: 5124-5131**.**

Del'Alamo L, d'Azevedo PA, Strob AJ, Rodríguez-Lopez DV, Monteiro J, Andrade SS, Pignatari AC, Gales AC. 2007. An outbreak of catalase-negative methicillin-resistant *Staphylococcus aureus.* J Hosp Infect 65:226-230.

Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci U S A 99:7687-7692.

Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt BG, Moore CE, and Day NP. 2003. How clonal is *Staphylococcus aureus*?. J Bacteriol 185: 3307–3316.

Foster TJ, and Hook M. 1998. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 6:484-488.

Gomes AR, Vinga S, Zavolan M, de Lencastre H. 2005. Analysis of the genetic variability of virulence-related loci in epidemic clones of methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 49:366-379.

Hartford O, Francois P, Vaudaux P, Foster TJ. 1997. The dipeptide repeat region of the fibrinogen binding protein (Clumping Factor) is required for the functional expression of the fibrinogen binding domain on the *Staphylococcus aureus* cell surface. Mol Microbiol 25:1065-1076.

Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabbinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci U S A 101:9786-9791.

Huynh HT, Robitaille G, Turner JD. 1991. Establishment of bovine mammary epithelial cells (MAC-T): an in vivo model for bovine lactation. *Exp Cell Res* 197:191-199.

Hunter P. 1990. Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol 28:1903-1905.

Josefsson E, Hartford O, O'Brien L, Patti JM, Foster T. 2001. Protection against experimental *Staphylococcus aureus* arthritis by vaccination with clumping factor A novel virulence determinant. J Infect Dis 184:1572–1580.

Josefsson E, Kubica M, Mydel P, Potempa J, Tarkowski A. 2008. In vivo sortase A and clumping factor A mRNA expression during *Staphylococcus aureus* infection. Microb Pathog 44:103-110.

Koreen L, Ramaswamy SV, Graviss EA, Naidich S, Musser JM, Kreiswirth BN. 2004. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol 42: 792–799.

Koreen L, Ramaswamy SV, Naidich S, Koreen IV, Graff GR, Graviss EA, Kreiswirth BN. 2005. Comparative sequencing of the serine-aspartate repeat-encoding region of the clumping factor B gene (clfB) for resolution within clonal groups of *Staphylococcus aureus*. *J Clin Microbiol* 43: 3985-3994.

Kuhn G, Francioli P, Blanc DS. 2006. Evidence for clonal evolution among highly polymorphic genes in methicillin-resistant *Staphylococcus aureus.* J Bacterial 188: 169–178.

Kuhn G, Francioli P, Blanc DS. 2007. Double-locus sequence typing using *clfB* and *spa,* a fast and simple method for epidemiological typing of methicillin-resistant Staphylococcus aureus. J Clin Microbiol 45:54-62.

Lammers A, Kruijt E, van de Kuijt C, Nuijten PJ, Smith HE. 2000. Identification of *Staphylococcus aureus* genes expressed during growth in milk: a useful model for selection of genes important in bovine mastitis?. Microbiology 146: 981–987.

Malachowa N, Sabat A, Gniadkowski M, Krzyszton-Russjan J, Empel J, Miedzobrodzki J, Kosowska-Shick K, Appelbaum PC, Hryniewicz W. 2005. Comparison of multiple locus variable number tandem repeat analysis with pulsed field gel electrophoresis, *spa* typing and multilocus sequence typing for clonal characterization of *Staphylococcus aureus* isolates. J Clin Microbiol 43:3095–3100.

Moreillon P, Entenza JM, Francioli P, McDevitt D, Foster TJ, François P, Vaudaux P. 1995. Role of *Staphylococcus aureus* coagulase and clumping factor in pathogenesis of experimental endocarditis. Infect Immun 63: 4738–4743.

Moxon R, Bayliss C, Hood D. 2006. Bacterial contingency loci: The role of simple sequence DNA repeats in bacterial adaptation. Annu Rev Genet 40:307-333.

Pericone CD, Bae D, Shchepetov M, McCool T, Weiser JN. 2002. Short-sequence tandem and nontandem DNA repeats and endogenous hydrogen peroxide production contribute to genetic instability of *Streptococcus pneumoniae*. J Bacteriol 184:4392-4399.

Park AM, Li Q, Nagata K, Tamura T, Shimono K, Sato EF, Inoue M. 2004. Oxygen tension regulates reactive oxygen generation and mutation of *Helicobacter pylori*. Free Radic Biol 36:1126-1133.

Qazi SN, Harrison SE, Self T, Williams P, Hill PJ. 2004. Real-time monitoring of intracellular *Staphylococcus aureus* replication. J Bacteriol 186:1065–1077.

Rabello RF, Moreira BM, Lopes RM, Teixeira LM, Riley LW, Castro AC. 2007. Multilocus sequence typing of *Staphylococcus aureus* isolates recovered from cows with mastitis in Brazilian dairy herds. J Med Microbiol 56:1505-1511.

Risley AL, Loughman A, Cywes-Bentley C, Foster TJ, Lee JC. 2007. Capsular polysaccharide masks clumping factor A–mediated adherence of *Staphylococcus aureus* to fibrinogen and platelets. J Infect Dis 196:919-927.

Roberson JR, Fox LK, Hancock DD, Gay JM, Besser TE. 1998. Sources of intramammary infections from *Staphylococcus aureus* in dairy heifers at first parturition. J Dairy Sci 81:687–693.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-386.

Sabat A, Krzyszton-Russjan J, Strzalka W, Filipek R, Kosowska K, Hryniewicz W, Travis J, Potempa J. 2003. New method for typing *Staphylococcus aureus* strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates. J Clin Microbiol 41:1801-1804.

Salaun L, Ayraud S Saunders NJ. 2005. Phase variation mediated niche adaptation during prolonged experimental murine infection with *Helicobacter pylori.* Microbiology 151:917–923.

Seegers H, Fourichon C, Beaudeau F. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. Vet Res 34:475-491.

van Belkum A. 2007. Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). *FEMS* Immunol Med Microbiol 49:22–27.

van Leeuwen WB, Melles DC, Alaidan A, Al-Ahdal M, Boelens HA, Snijders SV, Wertheim H, van Duijkeren E, Peeters JK, van der Spek PJ, Gorkink R, Simons G, Verbrugh HA, van Belkum A. 2005. Host- and tissue-specific pathogenic traits of *Staphylococcus aureus.* J Bacteriol 187:4584–4591.

Vernachio J, Bayer AS, Le T, Chai YL, Prater B, Schneider A, Ames B, Syribeys P, Robbins J, Patti JM. 2003. Anti-clumping factor A immunoglobulin reduces the duration of methicillin-resistant *Staphylococcus aureus* bacteremia in an experimental model of infective endocarditis. *Antimicrob Agents Chemother* 47:3400–3406.

Verstrepen KJ, Jansen A, Lewitter F, Fink GR. 2005. Intragenic tandem repeats generate functional variability. Nat Genet 37:986-990.

Wenzel RP, Edmond MB. 2001. The impact of hospital-acquired bloodstream infections. Emerg Infect Dis 7:174–177.

Wilson DJ, Gonzalez RN, Das HH. 1997. Bovine mastitis pathogens in New York and Pennsylvania: prevalence and effects on somatic cell count and milk production. J Dairy Sci 80:2592-2598.

Yamamura E, Nunoshiba T, Nohmi T, Yamamoto K. 2003. Hydrogen peroxide-induced microsatellite instability in the *Escherichia coli* K-12 endogenous tonB gene. Biochem Biophys Res Commun 306:570-576.

Zadoks RN, van Leeuwen WB, Kreft D, Fox LK, Barkema HW, Schukken YH, van Belkum A. 2002. Comparison of *Staphylococcus aureus* isolates from bovine and human skin, milking equipment, and bovine milk by phage typing, pulsed-field gel electrophoresis, and binary typing. J Clin Microbiol 40:3894-3902.

CONNECTING STATEMENT 1

In the previous study, development and evaluation of coding intragenic repeat-based sub-typing method was carried out using the hypervariable repeats in the *clfA*-R domain. Locus as well as repeat copy-number stability and suitability evaluation were vigorously assessed. Development of objective typing markers that are capable of coupling epidemiological data with colonization and pathogenic potential was crucial to identify host- and tissue-specialized lineages. Thus, the next chapter was designed (1) to investigate the correlation between the lengths of *clfA* (expressed in the number of copies of repeats in R domain, as the rest is conserved) and the host organs from where the isolates were isolated using human and intra-mammary isolates of different geographical origin, and (2) to examine the ability of the marker to recognize and group global populations of mammary-specific *S. aureus* from a pool of isolates.

CHAPTER III. ORGAN AND HOST-SPECIFIC CLONAL GROUPS OF STAPHYLOCOCCUS AUREUS FROM HUMAN INFECTIONS AND BOVINE MASTITIS REVEALED BY THE CLUMPING FACTOR A GENE

Kamaleldin B Said[1], Guoqiang Zhu[2] and Xin Zhao[1*]

[1]McGill University, Department of Animal Science, Ste. Anne de Bellevue, Quebec H9X 3V9, Canada

[2] College of Veterinary Science, Yanzhou University, China

[*]Corresponding author. Tel.: 1-514 398 7975; Fax: 1-514 398 7964.

Foodborne Path Dis. 2010. 7:111-119.

ABSTRACT

Staphylococcus aureus has become a major concern in public health due to the rapid evolution of resistant and host/organ specialized lineages adapted to humans and major food animals. However, the mechanism(s) of host- and organ-specialization in *S. aureus* is presently ill-defined. The objective of this study was to investigate whether coding intragenic repeat containing markers would be capable of detecting and grouping adaptive clones, tracing their sources, or studying the basis for their specializations in different microenvironments. We have analyzed the number of copies as well as the nucleotide sequences in the hypervariable R domain of the clumping factor A gene in 95 isolates from different organs in human patients and from bovine mastitis, using PCR, sequencing, multiple sequence alignment, and phylogenetic relationships. The results showed that isolates of the same human organ were polymorphic for *clfA*, while those of the mammary were clonal. Twenty out of the 23 sputum isolates had lower copy-numbers of 43-48, while 21 out of the 24 skin isolates had 55-63 copies. Twenty four repeat-types were identified with the index of discrimination of 0.9. Repeat-types and overall sequence pattern-groups were highly consistent. In addition, sequence alignments and phylogenetic analysis placed isolates from different hosts and organs into respective clusters. Thus, *clfA* is useful in detecting the clonal lineage of mastitis, tracing and grouping organ-specific strains, and might be a potential tool for studying host specialization and selection.

INTRODUCTION

Staphylococcus aureus (*S. aureus*) has emerged as an important public health concern globally. Direct cost due to methicillin-resistant *S. aureus* (MRSA) in Canada averaged $82 million in 2004 and could reach $129 million by 2010 (Goetghebeur et al., 2007). More MRSA infections have been reported to originate from community associated (CA)-MRSA (Klevens et al., 2006; Moran et al., 2006; Maree et al., 2007). Of particular importance is the recent emergence, essentially by diversification and clonal expansion, of the invasive CA-MRSA (USA300) (Kennedy et al., 2008; Klevens et al., 2007). In the dairy industry, *S. aureus* bovine mastitis is one of the most frequent and most costly disease (Seegers et al., 2003). It causes contagious mastitis, which usually develops into chronic intramammary infections resulting in poor success rates of antibiotic treatment (Gruet et al., 2001) due to their ability to specifically invade and persist in mammary cells (Hensen et al., 2000; Bayles et al., 1998; Qazi et al. 2004). However, the basis for emergence of host- and organ-specialized clones in mastitis and human (van Leeuwen et al., 2005; Kennedy et al., 2008) is quite elusive.

Hypervariable tandem-repeats have been useful in strain-typing and are potential tools for tracing and studying adaptations (Verstrepen et al., 2005). The rapid evolution of clones is potentially partly mediated by alternative strategies such as adaptive expressions in core genes (Li et al., 2009) and non-lethal mutations involving RecA independent mechanism of repeat variations in core-variable genes. It is well known that expression of repeat-containing genes that are in-frame for translation and do not alter the reading frame, code for the same protein but with different amino acid length leading to adaptive functional changes in the protein (Madoff et al., 1996; Citti et al., 1997; Simmons et al., 2004). For example, the Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMMs) (Foster and Hook, 1998) are encoded by the core variable genome and mostly contain tandem-repeats that could be responsible for adaptations (Verstrepen et al., 2005).

An important class of the MSCRAMMs is the fibrinogen-binding proteins ClfA and ClfB encoded by *clfA* and *clfB* genes, respectively (Ni Eidhin et al., 1998; McDevitt et al., 1997). They belong to a MSCRAMM subfamily called Sdr family (Ser-Asp

direpeat multigene family) (Josefsson et al., 1998). Repeat structure and modes of expression suggest their adaptive role. For instance, the copy number of repeat in the *clfA*, which influences the length of the whole encoded protein, affected the adherence and clumping titers of the cocci *in vitro* (Risley et al., 2007; Hartford et al., 1997). Furthermore, while ClfB activity is evident only in early growth phase and under aerobic conditions (Ni Edhin et al., 1998), ClfA is constitutively expressed and becomes abundant during deep-seated infections (Wolz et al., 1996, 2002; Josefsson et al., 2008). Thus, primary screening of strains with universal 84dhesion intragenic repeats such as *clfA*, may facilitate detection and grouping of strains with common host/organ-specific factors.

The R-domain of *clfA* is composed of 18bp repeats that follow the consensus [GAP (P= pyrimidine), TCN (N=any base), GAP, TCN, GAP, AGP] and code for [Ser-Asp]$_3$ dipeptide repeat (Foster and Hook, 1998). Repeat regions in *clfA* and *clfB* have been successfully applied for typing *S. aureus* (Kuhn et al., 2007; Sabat et al., 2003). The use of *clf*-R-domains in typing has also been postulated as useful additions to major molecular methods; particularly for detecting macro- and micro-variations in emerging clones in human infections (Gomes et al., 2005; Koreen et al., 2004; 2005). However, correlation of repeat variability to host and organ/tissue has not yet been investigated so far. This needs attention because specific tissue microenvironments are responsible for the differentiation of specialized clones (Josefsson et al., 2008). We have recently shown that the *clfA*-R domain was capable of identifying host-specific clones (Said et al., 2009). In the present report we aim to examine whether there was any correlation between the lengths of *clfA* (expressed in the number of copies of repeats in R domain, as the rest is conserved), and the host organs from where the isolates were isolated involving diverse population of patients, and from different dairy cows in different countries, and to explore whether it was possible to group host- or organ-specific clones using copy numbers as well as the sequence information.

MATERIAL AND METHOD

Bacterial strains

A total of 95 *S. aureus* strains/isolates were used in this study (Table 3.1). These included 23 isolates from skin abscesses of young people aged from two-year-old to 12-year-old that were collected between 2004 and 2007 and 24 isolates from human sputum. Skin isolates (n=24) were from The First People Hospital of Yangzhou which is a key regional health center in Jiangsu province and affiliated to Yangzhou University (kindly provided by Dr. Xiaofang Zhu). Isolates from sputum (n=23) were from the Chinese Disease Control (CDC) in Yanzhou, Jiangsu (kindly provided by Dr. Guoxiang Cao). This disease control center is independent from the hospital. The isolates were from 47 different patients.

Mastitis isolates were from sporadic cases in 37 different cows, 14 different herds that were located at remote regions in different countries. Chinese isolates (14), from 5 herds at different geographical locations, were collected and stored at the College of Veterinary Science, Yangzhou University. All were from milk samples taken for laboratory analysis of mastitis cases; each represented a single case (cow). Canadian isolates (22) were from 8 different herds, and each from a cow, were generously provided by Drs. Grant Tomita and Serge Messier; the Clinical Laboratory of Faculté de Médecine Vétérinarire, Université de Montreal, Quebec, Canada. In addition, eight sequences of isolates representing major lineages and virulence types (MRSA, CA-MRSA, CA-MSSA, and RF122 mastitis isolate) from different parts of the world were also included to further increase the diversities of regions (Table 3.1), in addition to three reference laboratory strains Smith, Lowenstein, and Wright. Glycerol stock cultures were stored at -80°C.

Table 3.1 The *S. aureus* strains/isolates and their source

Strains*	Site	Origin	Copy#
S.aureus aureus COL(MRSA)	http://www.tigr.org	Human	50.5
S.aureu MRSA252(HA-MRSA)	http://www.sanger.ac.uk/	Human	66
S.aureus N315 (HA-MRSA)	http://www.staphylococcus.org/jp/	Human	59.5
S. aureus MSSA476(CA-MRSA)	http://www.sanger.ac.uk/	Human	49
S.aureus MW2 (CA-MRSA)	http://www.bio.nite.go.jp/	Human	51
S.aureus. Mu50 (MRSA)	http://www.staphylococcus.org/jp/	Human	49.5
S. aureus NCTC 8325	http://microgen.ouhsc.edu/s_aureus/s_aureus_home.htm	Human	48
S. aureus USA300 (CA-MRSA)	http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj&cmd=Retrieve&dopt=Overview&list_uids=16313	Human	49
Mastitis *S. aureus* RF122	http://pathogenomics.umn.edu/staph_index.htm	Bovine	46

Isolate	**Host**	**Source**	**Isolate**	**Host**	**Sourc**	**Isolat**	**Host**	**Sourc**	**Isolat**	**Host**	**Sourc**	**Isolate**	**Host**	**Source**	**Isolate**	**Host**	**Source**
cdc 1	Human	China	cdc 22	Human	China	zdf 3	Human	China	zdf 22	Human	China	B35	Cow	China	Cq4	Cow	Canada
cdc 2	Human	China	cdc 23	Human	China	zdf 4	Human	China	zdf 23	Human	China	B36	Cow	China	Cq5	Cow	Canada
cdc 4	Human	China	cdc 24	Human	China	zdf 5	Human	China	zdf 24	Human	China	B37	Cow	China	Cq6	Cow	Canada
cdc 7	Human	China	cdc 25	Human	China	zdf 6	Human	China	Hz26	Human	China	B38	Cow	China	Cq7	Cow	Canada
cdc 8	Human	China	cdc 26	Human	China	zdf 7	Human	China	zdf 29	Human	China	B39	Cow	China	Cq8	Cow	Canada
cdc 9	Human	China	cdc 27	Human	China	zdf 9	Human	China	zdf 31	Human	China	B40	Cow	China	Cq9	Cow	Canada
cdc 10	Human	China				Hz10	Human	China	zdf 32	Human	China	B41	Cow	China	Cq10	Cow	Canada
cdc 11	Human	China				zdf 11	Human	China	zdf 33	Human	China	B42	Cow	China	Cq11	Cow	Canada
cdc 12	Human	China				zdf 12	Human	China				B43	Cow	China	Cq12	Cow	Canada
cdc 13	Human	China				zdf 13	Human	China				B44	Cow	China	Cq13	Cow	Canada
cdc 14	Human	China				zdf 14	Human	China				B45	Cow	China	Cq14	Cow	Canada
cdc 15	Human	China				zdf 15	Human	China				B46	Cow	China	Cq15	Cow	Canada
cdc 16	Human	China				zdf 16	Human	China				S04	Cow	China	Cq16	Cow	Canada
cdc 17	Human	China				zdf 17	Human	China				S14	Cow	China	Cq17	Cow	Canada
cdc 18	Human	China				zdf 18	Human	China							Cq18	Cow	Canada
cdc 19	Human	China				zdf 19	Human	China							Cq19	Cow	Canada
cdc 20	Human	China				zdf 21	Human	China							Cq20	Cow	Canada

[a]MRSA: methicillin resistant *Staphylococcus aureus*; HA-MRSA: hospital acquired MRSA; CA-MRSA: community associated MRSA. CA-MSSA: community associated methicillin sensitive *S.aureus.* Published strains were also used previously in Said et al. (2009). Sputum isolates are give code "cdc" followed by isolate numbers, and skin isolates as "zdf" followed by isolate numbers. For isolates, letters B and S associated to numbers indicate mastitis isolates from China, Cq stands for isolates from Canada, Quebec province.

Amplification of clfA R-domain in S. aureus isolates/strains

The Tandem Repeat Finder (TRF) program (http://tandem.bu.edu/trf/trf.html) (v 4.0) (Benson, 1999) was used to determine the number of copies of repeats in the R domain of *S. aureus* COL (50.5 copies) and the flanking sequences for primer design. This information was confirmed and used for designing PCR primers with the help of DNAstar and Primer3 (Rosen and Skaletsky, 2000) programs on conserved regions on the chromosome of *S. aureus* COL after aligning *clfA* sequences obtained from sequenced strains (Table 3.1). Sequences were BLAST checked for homologous regions. Primer sequences were designed with the forward primer as 5' TCCTGAACAACCTGATGAGC 3' and the reverse primer as 5' AGGTGAATTAGGCGGAACTAC 3'. Primers were synthesized by Shanghai GeneCore Biotechnologies Co. Ltd (Shanghai, China). PCR reactions were carried out with a Thermo Scientific Hybaid Px2 Thermal Cycler and the conditions that allowed amplification of single PCR products were established as initial denaturation at 94 ºC for 4 min followed by 25 cycles each of denaturation at 94 ºC for 1 min, annealing at 58 ºC for 1 min, extension at 72 ºC for 1 min, and a final extension of 72 ºC for 8 min. All PCR reactions were run in a 50 µl volume using reagents from TaKaRa, China. Each reaction contained 10 mM dNTP, 5 µl of 10× PCR buffer, 10 µM of each primer, 5 U μl^{-1} rTaq enzyme, and 60 ng DNA templates. Repetitive DNA are known to be unstable in general and to avoid extensive treatments a short and efficient protocol was used to isolate DNA from 18 h old cultures grown in tryptic soy broth (Sigma-Aldrich Ltd, Oakville, Ontario, Canada); this involved incubation of bacterial pellets for 30 min at 37ºC in lysis buffer (20 mM Tris-HCl pH 8.0, 2 mM sodium EDTA, 1.2% Triton X100 (Sigma), 20 mg ml^{-1} lysozyme (Sigma), and 10 µg ml^{-1} lysostaphin (Sigma), followed by digestion with proteinase K (Invitrogen Canada Inc., Burlington,

Ontario, Canada).at 55ºC for 30 min. Then the lysates were boiled for 5 minutes and centrifuged at 8000rpm for 1 minute. The supernatants were carefully collected checked for quality and quantity by gel electrophoresis and spectrophotometers before using as templates.

Fragments were separated on 1% agarose gel using 1% TBE as running and gel buffer in BioRad DNA Sub-Cell Electrophoresis Tank (dimensions; 12′, 6′, and 3′, Voltage 100, runtime 2.5h). Images were taken on gel photo doc Jieda801 (Jiangsu Jieda Technology Co. Ltd., China). Differences in R-domain copy numbers were obtained from PCR product sizes on gels using information on *S. aureus* COL. The copy numbers were confirmed by sequencing PCR products at the Shanghai GeneCore Biotechnologies Co. Ltd (Shanghai, China). Variant types in this study were assigned based on difference in copy numbers (repeat-types); a difference by a single (18bp) copy was considered a repeat-type.

Nucleotide sequence analysis and phylogenetic relationships

Chromatograms and text files of the resulting sequences were checked and assembled using Chromas v2.32 program. Copy numbers of repeats on the sequences were checked with the help of the TRF program and confirmed by visual sorting of individual copy numbers. Then, sequences of individual strains and isolates were analyzed to confirm copy numbers using Chromas, and ClustalW from European Biofinformatics Institute sites (http://www.ebi.ac.uk/Tools/clustalw/index.html). Afterwards, global multiple sequence alignments of the strains and isolates were carried out using Geneious Bioinformatics package v3.8 to reveal the nucleotide sequence similarity patterns (sequence groups) of related strains. Since the repeats in the R-domain are highly degenerate and the sequences vary considerably at both 3’ and 5’ ends of the R region, comparative phylogenetic grouping was carried out using whole R-domain sequences including parts of highly conserved regions with the help of the Neighbor-Joining method built in the Geneious bioinformatics package (v3.8). The cladograms were based on the most likely groupings that reflect genetic relationships of strains and involved circular trees.

Discriminatory power

The Discriminatory Power Calculator available the following internet site (http://biophp.org/stats/discriminatory_power/demo.php) was used to calculate the index of discrimination (ID); [ID = 1] indicates the ability to differentiate each isolate, and [ID = 0] indicates that all isolates are identical. Determination of repeat-types was based on differences in the number of the 18-bp copies of tandem repeats contained in the R-domain of the *clfA* gene, and confirmed by sequence. The numerical index of discriminatory power was used to give numeric estimates for strain differentiation; and the values (defined as the average probability that the typing system will assign a different type to two unrelated strains randomly sampled in the microbial population of a given taxon) were estimated according to Hunter (1990). The value of 1 indicates the ability to differentiate each isolate, and 0 indicates that all isolates are identical.

RESULTS

Amplified clfA R-domain reveals organ-specific groups of S. aureus repeat-types

In order to explore whether *S. aureus* would display organ-specific *clfA* length, we have amplified the hypervariable R domain, from isolates of different sites of human patients and from bovine mastitis from different dairy herds located either in China or in Canada. Figure 3.1 shows differences in R-domain copy numbers, of representative number of isolates, obtained from PCR product sizes on gels. As shown in Table 3.2, the human sputum (n=23) and skin (n= 24) isolates had R domain copy numbers in the narrow range of 43-48 (except three) and 55-63 (except four), respectively. Two published HA-MRSA strains (MRSA252, N315; the former is a dominant hospital acquired MRSA in the UK, and the latter is a MRSA strain isolated in 1982 in Japan) had higher copy numbers (66 and 59.5 copies, respectively) than the low copy number group (43 to 52 copies) consisting all other published human strains (n=6) including all 4 community acquired strains as well as 37 mastitis strains/isolates (Table 3.2). Twenty eight out of 37 mastitis isolates analyzed (76%) from the two countries had 52 copies of repeats.

The repeat profiles based on 18bp copy differences were uniquely organ-specific (Table 3.2). The sputum isolates were of five *clfA* repeat-types namely, T, B, D, C, and E; the vast majority of them (20/23) belonged to types T and D with copy numbers of 43 and 48, respectively. Repeat-types B, C, and E comprised of a single isolate each. Similarly, the skin isolates formed eight different types F, G, H, R, J, K, L, and M. However, except for F (copy number of 58), all contained less than five isolates each, and isolate zdf29 shared T repeat-type with skin isolates. The 37 mastitis isolates, (fourteen mastitis isolates from China and 22 from Canada (17 new isolates plus 5 previously identified clones) belonged to two common repeat-types found in both countries; a major type X had 28 isolates (X_{cn} for China type and X_{cq} for Quebec, Canada), and a minor type Y contained only four isolates (Y_{cn} China isolates and Y_{cq} for Canada). In addition, a third type Z (isolates Cq10, 760 and 855 with 50 copies), was found only in Canada; while, a fourth type S04 (47 copies) was unique to China. Further, a fifth type Q contained only the RF122 lineage. In this study, 19 repeat-types were identified only in human strains tested with an ID of 0.9; and overall 24 types were identified with an ID of 0.9 (Table 3.2).

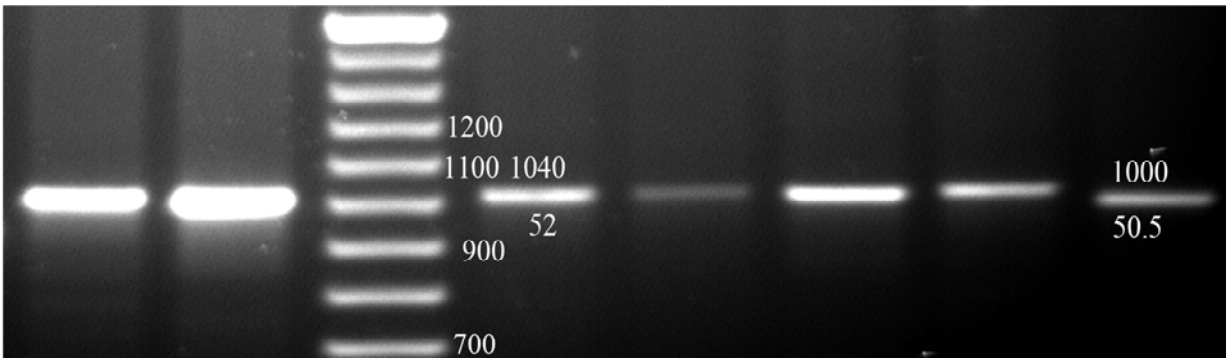


Figure 3.1 Amplified whole R domain length polymorphism (numbers above the bands are R domain sizes) and the corresponding repeat variations (numbers below the bands are the corresponding copy numbers) in representative isolates.

Table 3.2: Genotypic profiles of all isolates and strains. The repeat-types were based on copy numbers of the R domain. The repeat-type letters Y and X, were given to the same types found in China and Canada.

CDC isolates from sputum			**Zdf isolates from skin abscesses**			**Published strains**			**Mastitis isolates China**			**Mastitis isolates Canada**		
I[a]	C[b]	RT[c]	I	C	RT	S[d]	C	RT	I	C	RT	I	C	RT
cdc1	48	$T_{cdc\ 1,10,14,24,25,26,27,zdf29,NCTC8325}$	zdf3	58	$F_{zdf\ 3,6,10,14,16,17,22,32}$	Mu50	49.5	P	B35	52	X_{cn}	Cq4	52	X_{cq}
cdc2	53	$B_{cdc\ 2}$	zdf4	60	$G_{zdf\ 4,11,26}$	RF122	46	Q	B36	52	X_{cn}	Cq5	52	X_{cq}
			zdf 5	31	$H_{zdf\ 5}$	COL	50.5	R	B37	44	Y_{cn}	Cq6	52	X_{cq}
cdc4	43	$D_{cdc\ 4,7,8,9,11,13,16,17,18,19,20,22,23}$	zdf6	58		USA300	49	S	B38	52	X_{cn}	Cq7	52	X_{cq}
cdc7	43		zdf7	50.5	$R_{zdf\ 7}$	NCTC8325	48	T	B39	52	X_{cn}	Cq8	52	X_{cq}
cdc8	43		zdf9	55.5	$J_{zdf\ 9,12,13,23}$	MW2	51	U	B40	52	X_{cn}	Cq9	52	X_{cq}
cdc9	43		zdf10	58		MSSA476	49	S	B41	52	X_{cn}	Cq10	50.5	Z_{cq}
cdc10	48		zdf11	60		Lowenstein	50	C	B42	52	X_{cn}	Cq11	44	Y_{cq}
cdc11	43		zdf12	55.5					B43	52	X_{cn}	Cq12	52	X_{cq}
cdc12	50	$C_{cdc\ 12}$	zdf13	55.5		Smith	66	V	B44	52	X_{cn}	Cq13	52	X_{cq}
cdc13	43		zdf14	58					B45	52	X_{cn}	Cq14	52	X_{cq}
cdc14	48		zdf15	65.5	$K_{zdf\ 15}$				B46	52	X_{cn}	Cq15	52	X_{cq}
cdc15	55	$E_{cdc\ 15}$	zdf16	58		Wright	57	I	S04	47	$S04_{cn}$	Cq16	52	X_{cq}
cdc16	43		zdf17	58		MRSA252	66	V	S14	44	Y_{cn}	Cq17	52	X_{cq}
cdc17	43		zdf18	63	$L_{zdf\ 18,24,31,33}$	N315	59.5	W				Cq18	52	X_{cq}
cdc18	43		zdf21	52	$M_{zdf\ 21}$							Cq19	52	X_{cq}
cdc19	43		zdf22	58								Cq20	52	X_{cq}
cdc20	43		zdf23	55.5										
cdc22	43		zdf24	63										
cdc23	43		zdf26	60										
cdc24	48		zdf29	48	T									
cdc25	48		zdf31	63										
cdc26	48		zdf32	58										
cdc27	48		zdf33	63										

[a,b,c,d] Subheadings I, C, RT, and S, stand for Isolates, Copy numbers, Repeat-Types, and Strains, respectively. Repeat-Types were designated by letter and the number of isolates associated to each type was given in lower case numerals associate to strain codes. (e.g., $T_{cdc\ 1,10,14,24,25,26,27,zdf29,NCTC8325}$).

Multiple nucleotide sequence alignment patterns of isolates

Multiple nucleotide sequence analysis was carried out using whole R domain nucleotide sequences to examine alignment patterns of isolates from different organs, and to validate the groupings based on copy numbers. As shown in Figure 3.2, three major sequence-groups (I, II and III) can be easily seen from the output. Group I contained bovine mastitis-associated *S. aureus* that were separated into two distinct but very closely related pattern-types. A globally dominant pattern group contained almost all isolates from China (S04, S14, B35, B37, and B38 to B46), and 19 isolates from Canada (4m, 760, 855, and 764(16). Another minor but very closely related group contained the sequenced mastitis RF122 isolate as well as 140 and B37 from Canada and China, respectively. A notable difference between the patterns of the two mastitis groups is the repeat-unit deletions in the center of the minor-group sequences (nucleotide (nt) 690-750). Only one isolate (B36) had a uniquely distinct sequence-type from that of mastitis group; and although it exhibited the same mastitis-minor group deletion as above, sequence similarity to human isolates placed it as an independent lineage more closely related to those from the skin.

Group II, contained exclusively skin isolates. Despite the high-level sequence similarities, extended deletions in nt 690-750 and downstream regions created four subtypes and four independent lineages within this group. Group III was predominantly sputum isolates and published CA-MRSAs strains. This group also shared the central deletion found in the mastitis-minor group. In addition, there were three main deletions (nt 60-100, 500-600, and 900-1000) that created two subtypes within this group; one exclusive for cdc (sputum) isolates and the other mixture of sputum and invasive CA-MRSAs. In addition, a common sequence region above GroupI was shared by sputum isolates zdf09, zdf12, zdf13, zdf14, and zdf29, skins isolates cdc02, cdc12, cdc15, the reference strains Smith and Wright, as well as by the hospital strains MRSA252, N315,

and Mu50, which were mostly shown as independent lineages. Overall, members of each group showed similar nucleotide sequence patterns primarily differing only by the repeat-unit (copy-numbers) variations found mostly in the center and towards the ends. Worthy of mentioning is the similarity between HA-MRSAs N13 and MRSA252 that had the lowest repeat-deletions in overall strains analyzed, followed by the two CA-strains, MSSA476 and MW2 which differed from each other mainly by two deletions; one at nucleotide 189 and the other between nucleotide 300-400.

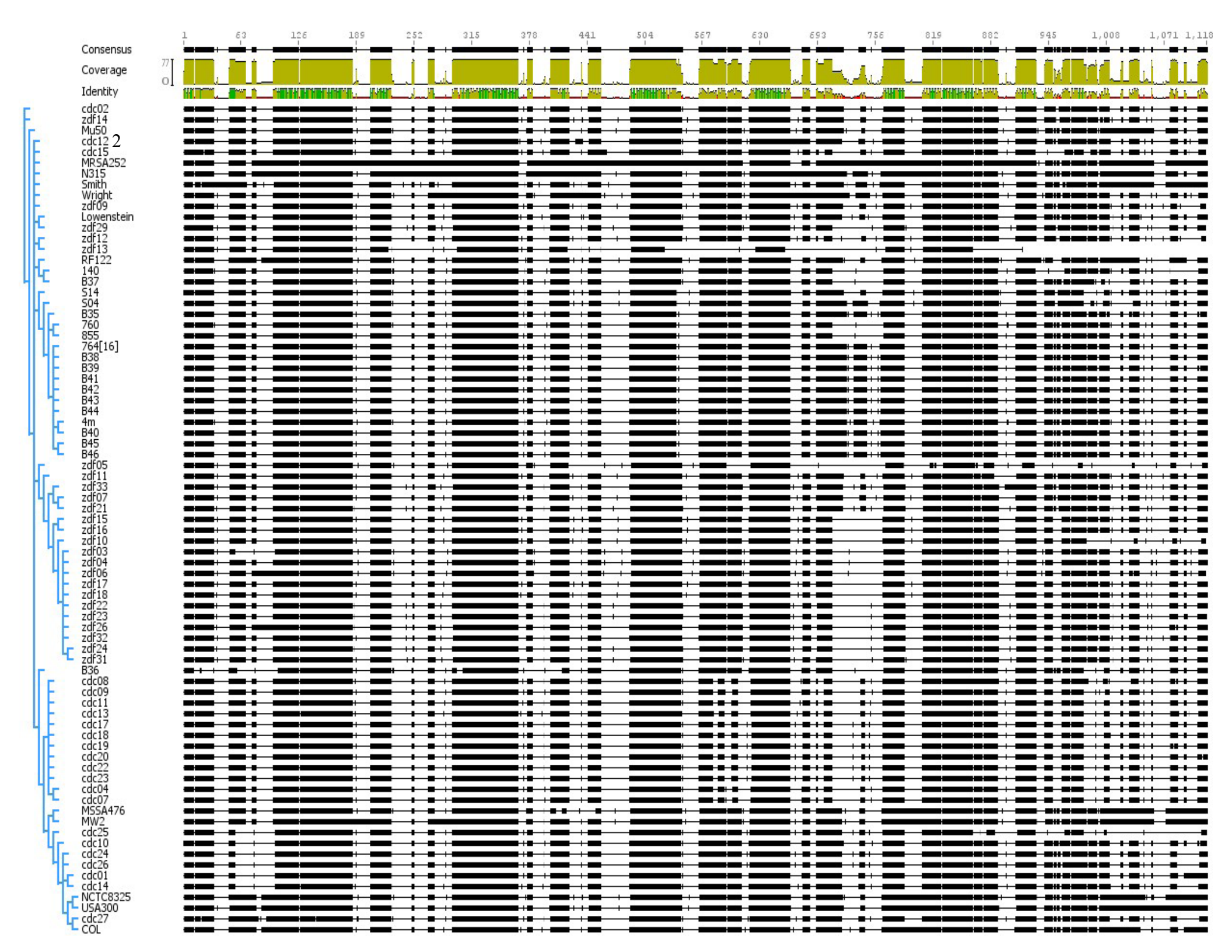

1
63
126
189
252
315
378
441
504
567
630
693
756
819
882
945
1,008
1,071
1,118
Consensus
Coverage
77
0
Identity
cdc02
zdf14
Mu50
cdc12
2
cdc15
MRSA252
N315
Smith
Wright
zdf09
Lowenstein
zdf29
zdf12
zdf13
RF122
140
B37
S14
S04
B35
760
855
764[16]
B38
B39
B41
B42
B43
B44
4m
B40
B45
B46
zdf05
zdf11
zdf33
zdf07
zdf21
zdf15
zdf16
zdf10
zdf03
zdf04
zdf06
zdf17
zdf18
zdf22
zdf23
zdf26
zdf32
zdf24
zdf31
B36
cdc08
cdc09
cdc11
cdc13
cdc17
cdc18
cdc19
cdc20
cdc22
cdc23
cdc04
cdc07
MSSA476
MW2
cdc25
cdc10
cdc24
cdc26
cdc01
cdc14
NCTC8325
USA300
cdc27
COL

Figure 3.2. Global sequence alignment patterns: using a combined guide tree based on Neighbor-Joining (Geneious Bioinformatics Package v3.8) involving the R-domain nucleotide sequences of 77 sequences. Seventeen mastitis isolates from Canada and cdc16 were not sequenced and thus were not included in this analysis. Sequence similarity and coverage are indicated by colors. For published sequences, standard abbreviations were used. For isolates, letters B and S associated to numbers indicate mastitis isolates from China. Isolates 4m, 140, 760, 764(16) and 855 were the previous mastitis isolates from Canada. Roman numerals I, II, and III indicate the main groups of isolates

Phylogenetic relationships of organ-specific groups

Because the repeats in the R domain sequence are irregular in the beginning and towards the ends and perfect only in the center, we thought that the genetic relatedness between closely related strains would best be examined using the whole sequence including parts of conserved 3' and 5 regions. Phylogenetic analysis of sequences is shown in Figure 3.3. Similar to the groupings shown by the copy number and nucleotide sequence alignments in Table 3.2 and Figure 3.2, Figure 3.3 also confirmed the overall genetic relationships of all strains tested. A colored tree represented each group. Except for five, isolates from human skin abscess were grouped together in a distinct cluster as shown by the green branch. Similarly, isolates from the sputum were included in a major cluster (brown) with two subclusters one of which contained the published CA-MRSAs (USA300, MW2 and MSSA476). All of the mastitis isolates tested (blue clusters), except for the Chinese isolate B36, were shown as closely related groups. Moreover, inclusion of 5 isolate sequences representing the previously identified Canadian clones (764, 4m, 760, 855, and 140) in strain pool as controls has resulted in their designation to correct clusters within the mastitis group (dark blue). In addition, the same independent lineages shown by the sequence pattern were confirmed by phylogenetic relationships.

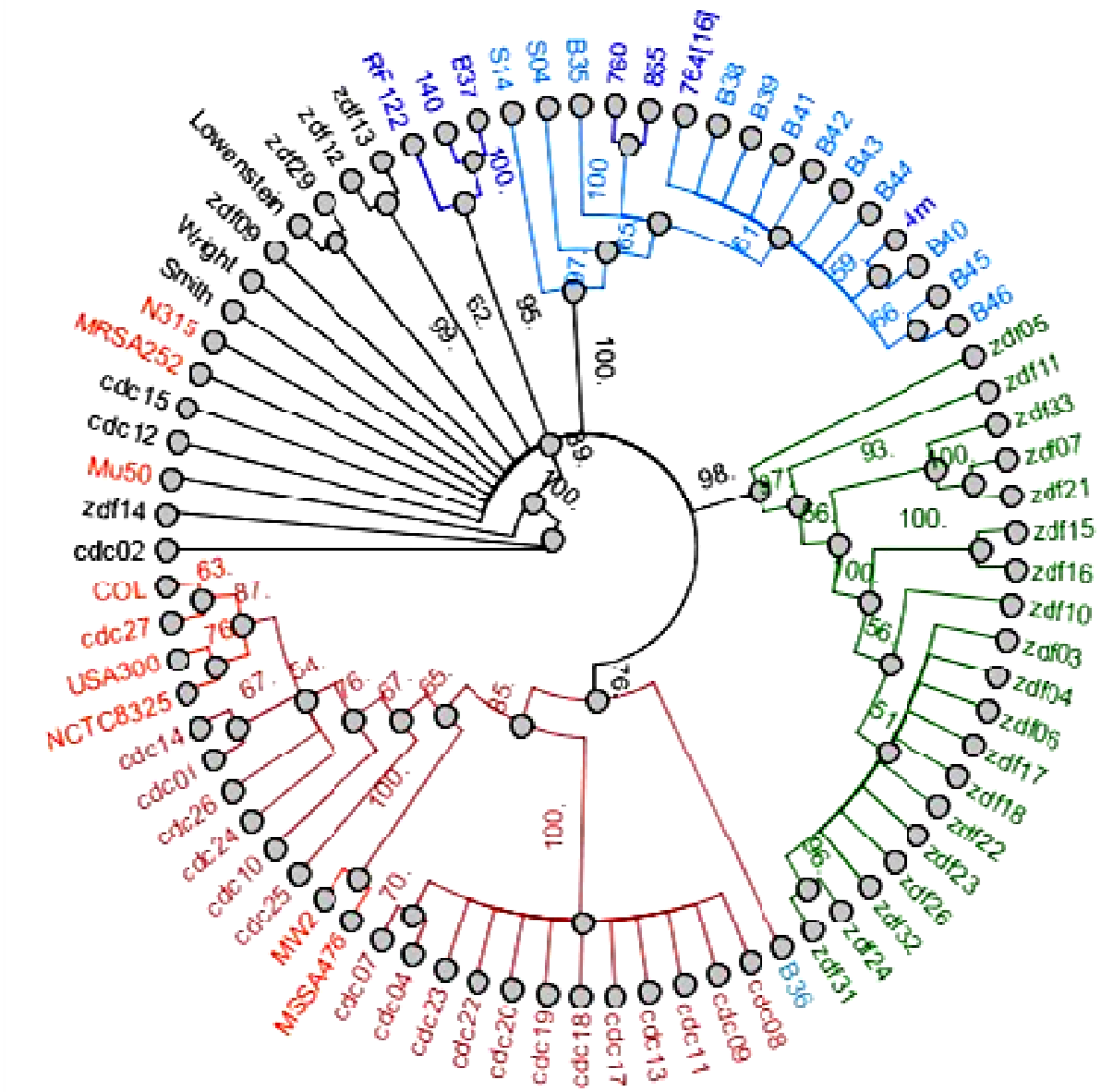


Figure 3.3. The phylogenetic relationships of isolates: circular tree based on the Neighbor-Joining program in Geneious bioinformatics package (v. 3.8). Distances were estimated by the Tamura–Nei model, bootstrapping100, threshold of 50%, using the R-domain sequences including conserved 3' and 5' regions. Colored trees indicate organ specific groups; the light blue are mastitis isolates from China, and the deep blue are those from Canada. The green branch is for the skin isolates, and the brown for sputum "cdc series". The red branches are the known published CA-MRSA, CA-MSSA, and MRSA strains.

DISCUSSION

We have recently evaluated the stability and use of *clfA* marker in typing mastitis and human isolates (Said et al., 2009). In the present study, we have proven the polymorphic nature of *clfA* from isolates of the same human organs, and the uniquely global clonal profiles of mammary isolates of different regions implying a potential selection in the mammary gland. Thus, after a precise evaluation of the *clfA* of clinical isolates of *S. aureus*, we provide evidence of, for the first-time to the best of our knowledge, a correlation between the origin of *S. aureus* (hosts/sites of infection) and the lengths of the repeated moiety of the target gene. This might be linked to virulence to various lineages of *S. aureus*; and therefore, constitutes a potential strategy for studying host-specialization and selection in *S. aureus* lineages.

Determination of copy number (Figure 3.1, Table 3.2) showed that the sputum isolates had low R domain copy numbers in the narrow range of 43 to 48 copies, except for three isolates. On the other hand, 21 of the 24 isolates from skin infections had a higher copy numbers in the range of 55 to 63. Interestingly, with the exception of the two published HA-MRSAs (MRSA252 dominant in UK, N315 isolate from Japan), all of published CA-MRSA strains as well as the mastitis isolates and almost all sputum isolates were in the uniquely low copy numbers narrow range of 43 to 52 copies, with bovine isolates at the upper end. Considering the various ranges of repeats for isolates from different organs, it appears that the variations in repeat copy numbers may be selected for adherence and invasion.

The copy-number based groupings were further supported by sequence information. Figures 3.2 and 3.3 showed that the published CA-MRSAs and CA-MSSA strains were subgrouped together and were closely related to isolates of the sputum (Group III). This shows that this group has genetic relationships based on *clfA*, and might be related to the fact that CA-MRSA and CA-MSSA strains have a predilection strategy to soft tissues such as the upper respiratory tract organs. CA-MRSA is one of the major causes of pneumonic bacteremia reaching rates as high as 62% (Rubinstein et al., 2008). Similar to the phylogenetic relationships of MRSA in Figure 3.3, a recent study also revealed the close relationships among staphylococcal strains COL, NCTC8325, and

USA300 and the relative evolutionary distance to strains MSSA476 and MW2 (Baba et al., 2008). Furthermore, using the repeat based system, we report herein that two HA-MRSAs (MRSA252 and N315) were found as related lineages, but distant from CA-MRSA and CA-MSSA strains, confirming the notion that the genetic backgrounds of CA-MRSA and HA-MRSA strains were uniquely different, and that the former lineage is distinct and descends from compatible endemic methicillin-susceptible *S. aureus* (MSSA), and not from the hospital-acquired (HA-MRSA) lineages (Groom et al., 2001; Naimi et al., 2001; Vandenesch et al., 2003).

Consistent with the findings reported using major molecular methods (van Leeuwen et al., 2005; Smith et al. 2005; Kapur et al., 1995), we were able to reach a similar conclusion that mastitis is caused by a limited number of global clones. However, our results were based on more specified organ/tissue specific isolates, and used intragenic repeats. The mastitis clusters supported by a 100% of consensus shown in Figure 3.3 included all of the Canadian and Chinese clones, except for one Chinese (B37) and one Canadian (140) lineages which were found identical and together with the known mastitis isolate RF122 formed an independent, but closely related mastitis group. Similarly, the minor Canadian type Z (comprising isolates 760 and 855), and the Chinese isolates S04, S14, and B37 formed distinct subclusters within the mastitis group, consistent with the finding of Smith et al., (2005) that diversifications into subclonal populations occurred in different geographical regions.

Comparison of repeat-types within the sequence groups I, II, and III showed a high degree of consistency demonstrating sequence conservation in isolates with similar repeat copy numbers. As shown in Table 3.2, sputum isolates had a narrow range of repeat copies 43-48 (except cdc02, cdc12, and cdc15 with higher copies 53, 50, and 55, respectively). Sequence pattern in Figure 3.2 grouped these in group III, except for the last three isolates which were more related to high-copy lineages similar to HA-strain MRSA252. Similarly, skin isolates were in the range of 55-63 copies and formed sequence group II except for four isolates (zdf12, zdf13, zdf14, and zdf29; with 55.5, 55.5, 58, and 48 copies, respectively, which were more closer to the three divergent lineages from the sputum. Mastitis isolates from different countries and regions had five repeat-types; a predominant clone X had 52 copies, while Y, Z, Q, and SO4, had 44, 50,

46, and 47, respectively. All mastitis isolates were placed in two very closely related subclusters within group I (Figure 3.2). Their overall sequence pattern was conserved and the grouping was consistent with that made by repeats. These suggest that the proposed repeat typing system is potentially useful for typing/grouping host- and organ- specific lineages.

In summary, we have provided new perspectives on the use of the *clfA* marker as a reproducible, rapid, sensitive, and cost-effective screening tool for organ-specific grouping of specialized clones from different host/environment, before application of major methods such as Multilocus Sequence Typing. Although the entire successful adaptational strategies of the species cannot be result of a single locus, our data suggest that repeat-mediated adaptational variations in the structure of this universal adhesin may serve as a potential mechanism for host/tissue specialization in *S. aureus.*

ACKNOWLEDGEMENT

This study was supported by funding from a discovery grant from Natural Science and Engineering Research Council of Canada and from the Canadian Bovine Mastitis Research Network. Additional support was provided by a grant from the National Nature Science Foundation of China (NSFC Grant No. 30828026). We would like to acknowledge Ms. J. He for her assistance in the laboratory.

REFERENCES

Baba T, Bae T, Schneewind O, Takeuchi F, Hiramatsu K. 2008. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of Staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. J Bacteriol 190:300-310.

Bayles KW, Wesson CA, Liou LE, Fox LK, Bohach GA, Trumble WR. 1998. Intracellular *Staphylococcus aureus* escapes the endosome and induces apoptosis in epithelial cells. Infect Immun 66: 336-342.

Benson G. 1999. Tandem Repeats Finder: a program to analyze DNA sequences. Nucleic Acids Res 27: 573-580.

Citti C, Kim MF, Wise KS. 1997. Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth-inhibiting host antibodies. Infect Immun 65:1773–1785.

Foster TJ and Hook M. 1998. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 6: 484-488.

Gomes AR, Vinga S, Zavolan M, de Lencastre H. 2005. Analysis of the genetic variability of virulence-related loci in epidemic clones of methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 49:366-379.

Groom AV, Wolsey DH, Naimi TS, Smith K, Johnson S, Boxrud D, Moore KA, Cheek, JE. 2001. Community-acquired methicillin-resistant *Staphylococcus aureus* in a rural American Indian community. JAMA 286:1201–1205.

Gruet P, Maincent P, Berthelot X, Kaltsatos V. 2001. Bovine mastitis and intramammary drug delivery: review and perspectives. Adv Drug Deliv Rev 50:245–259.

Goetghebeur M, Landry PA, Han D, Vicente C. 2007. Methicillin-resistant *Staphylococcus aureus*: A public health issue with economic consequences. Can J Infect Dis Med Microbiol 18:27-34.

Hartford O, Francois P, Vaudaux P, Foster TJ. 1997. The dipeptide repeat region of the fibrinogen binding protein (Clumping Factor) is required for the functional expression of the fibrinogen binding domain on the *Staphylococcus aureus* cell surface. Mol Microbiol 25:1065-1076.

Hensen SM, Pavicic MJ, Lohuis JA, Poutrel B. 2000. Use of bovine primary mammary epithelial cells for the comparison of adherence and invasion ability of *Staphylococcus aureus* strains. J Dairy Sci 83:418-429.

Hunter PR. 1990. Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol 28:1903-1905.

Josefsson E, Kubica M, Mydel P, Potempa J, Tarkowski A. 2008. *In vivo* sortase A and clumping factor A mRNA expression during *Staphylococcus aureus* infection. Microb Pathog 44:103-110.

Josefsson E, McCrea KW, Nı Eidhin D, O'Connell D, Cox J, Hook M, Foster TJ. 1998. Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus.* Microbiology 144:3387–3395.

Kapur V, Sischo WM, Greer RS, Whittam TS, Musser JM. 1995. Molecular population genetic analysis of *Staphylococcus aureus* recovered from cows. J Clin Microbiol; 33:376–380.

Kennedy AD, Otto M, Braughton KR, Whitney AR, Chen L, Mathema B, Mediavilla JR, Byrne KA, Parkins LD, Tenover FC, Kreiswirth BN, Musser JM, DeLeo FR. 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: Recent clonal expansion and diversification. Proc Natl Acad Sci U S A 105:1327-1332.

Klevens RM, Morrison MA, Fridkin SK, Reingold A, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Fosheim G, McDougal LK, Tenover FC. 2006. Active Bacterial Core Surveillance of the Emerging Infections Program Network. Community-associated methicillin-resistant *Staphylococcus aureus* and healthcare risk factors. Emerg Infect Dis 12:1991–1993.

Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Zell ER, Fosheim GE, McDougal LK, Carey RB, Fridkin SK. 2007. Active Bacterial Core surveillance (ABCs) MRSA Investigators. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. JAMA 298:1763-1771.

Koreen L, Ramaswamy SV, Graviss EA, Naidich S, Musser JM, Kreiswirth BN. 2004. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol 42: 792–799.

Koreen L, Ramaswamy SV, Naidich S, Koreen IV, Graff GR, Graviss EA, Kreiswirth BN. 2005. Comparative sequencing of the serine-aspartate repeat-encoding region of the clumping factor B gene (clfB) for resolution within clonal groups of *Staphylococcus aureus*. J Clin Microbiol 43: 3985-3994.

Kuhn G, Francioli P, Blanc DS. 2007. Double-locus sequence typing using *clfB* and *spa,* a fast and simple method for epidemiological typing of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 45:54-62.

Li M, Diep BA, Villaruz AE, Braughton KR, Jiang X, DeLeo FR, Chambers HF, Lu Y, Otto M. 2009. Evolution of virulence in epidemic community-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci U S A 106:5883-5888.

Madoff LC, Michel JL, Gong EW, Kling DE, Kasper DL. 1996. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc Natl Acad Sci USA 93: 4131-4136.

Moran, GJ, Krishnadasan A, Gorwitz RJ, Fosheim GE, McDougal LK, Carey RB, Talan DA. 2006. EMERGEncy ID Net Study Group. Methicillin-resistant *S. aureus* infections among patients in the emergency department. N Engl J Med 355:666–674.

Maree, CL, Daum RS, Boyle-Vavra S, Matayoshi K, Miller LG. 2007. Community-associated methicillin-resistant *Staphylococcus aureus* isolates causing healthcare-associated infections, Emerg Infect Dis 13:236–242.

McDevitt D, Nanavaty T, House-Pompeo K, Bell E, Turner N, McIntire L, Foster T, Hook M. 1997. Characterization of the interaction between the *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. Eur J Biochem 247:416-24.

Naimi TS, LeDell KH, Boxrud DJ, Groom AV, Steward CD, Johnson SK, Besser JM, O'Boyle C, Danila RN, Cheek JE, Osterholm MT, Moore KA, Smith KE. 2001. Epidemiology and clonality of community-acquired methicillin- resistant *Staphylococcus aureus* in Minnesota. 1996–1998. *Clin Infect Dis* 33:990–996.

Nı Eidhin, D, Perkins S, Francois P, Vaudaux P, Hook M, Foster TJ. 1998. Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. Mol. Microbiol 30: 245–257.

Qazi, SN, Harrison SE, Self T, Williams P, Hill PJ. 2004. Real-time monitoring of intracellular *Staphylococcus aureus* replication. J Bacteriol 186:1065–1077.

Risley AL, Loughman A, Cywes-Bentley C, Foster TJ, Lee JC. 2007. Capsular polysaccharide masks clumping factor A-mediated adherence of *Staphylococcus aureus* to fibrinogen and platelets. J Infect Dis 196:919–927.

Rozen S, and Skaletsky H J. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-386.

Rubinstein E. 2008. *Staphylococcus aureus* bacteraemia with known sources. Int J Antimicrob Agents 32: Suppl 1:S18-20.

Said, KB, Ramotar K, Zhu G, Zhao X. 2009. Repeat-based subtyping and grouping of *Staphylococcus aureus* from human infections and bovine mastitis using the R-domain of the clumping factor A gene. Diagn Microbiol Infect Dis 63:24–37.

Sabat A, Krzyszton-Russjan J, Strzalka W, Filipek R, Kosowska K, Hryniewicz W, Travis J, Potempa J. 2003. New method for typing *Staphylococcus aureus* strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates. J Clin Microbiol 41:1801-1804.

Seegers H, Fourichon C, Beaudeau F. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. Vet Res 34:457-91.

Simmons WL, Denison AM, Dybvig K. 2004. Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: Shield Hypothesis. Infect Immun 72: 6846–6851.

Smith EM, Green LE, Medley GF, Bird HE, Fox LK, Schukken YH, Kruze JV, Bradley AJ, Zadoks RN, Dowson CG. 2005. Multilocus sequence typing of intercontinental bovine *Staphylococcus aureus* isolates. J Clin Microbiol 3:4737-4743.

van Leeuwen WB, Melles DC, Alaidan A, Al-Ahdal M, Boelens HA, Snijders SV,Wertheim H, van Duijkeren E, Peeters JK, van der Spek PJ, Gorkink R, Simons G, Verbrugh HA, van Belkum A. 2005. Host- and tissue-specific pathogenic traits of *Staphylococcus aureus.* J Bacteriol 187: 4584–4591.

Vandenesch F, Naimi T, Enright MC, Lina G, Nimmo GR, Heffernan H, Liassine N, Bes M, Greenland T, Reverdy ME, Etienne J. 2003. Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. Emerg Infect Dis 9: 978-984.

Verstrepen KJ, Jansen A, Lewitter F, Fink GR. 2005. Intragenic tandem repeats generate functional variability. Nat Genet 37:986-990.

Wolz, C, Mcdevitt D, Foster TJ, Cheung AL. 1996. Influence of *agr* on fibrinogen binding in *Staphylococcus aureus* Newman. Infect Immun 64:3142–3147.

Wolz, C, Goerke C, Landmann R, Zimmerli W, Fluckiger U.2002. Transcription of clumping factor A in attached and unattached *Staphylococcus aureus* in vitro and during device-related infection. Infect. Immun 70:2758–2762.

CONNECTING STATEMENT 2

In Chapter II, we have developed and evaluated the suitability of the *clfA* marker as sub-typing molecular method, and in Chapter III we have examined the global clonal nature of mastitis *S. aureus* and the predilection of strains to different organs from where they were isolated. In the following chapter, we have carried out a comprehensive comparative analysis to rapidly track the distribution of dominant repeat types in different geographic location in Canada, to confirm strain relationships obtained by the *clfA* copy-number and sequence information as well as by using the gold standard methods, the Pulsed Field Gel Electrophoresis, and *spa*-typing.

CHAPTER IV. REGIONAL PROFILING FOR GENOTYPE DIVERSITY OF MASTITIS-SPECIFIC STAPHYLOCOCCUS AUREUS LINEAGE IN CANADA USING CLUMPING FACTOR A, PULSED FIELD GEL ELECTROPHORESIS, AND SPA-TYPING

Kamaleldin B Said[1], Johanne Ismail[2], Jennifer Campbell[3], Michael R. Mulvey[3], Anne-Marie Bourgault[2], Serge Messier[4], and Xin Zhao[1]*

[1]McGill University, Montreal, Quebec, Canada

[2]Laboratoire de santé publique du Québec, Institut National de santé publique du Québec, Montreal, Quebec, Canada.

[3]Canadian Nosocomial Surveillance Program, National Microbiology Laboratory, Public Health Agency of Canada, Health Canada, Winnipeg, Manitoba

[4]Faculté *de* médecine vétérinaire, Université de Montréal, St-Hyacinthe, Quebec, Canada.

[1]*Xin Zhao, corresponding author. Mailing address:Department of Animal Science**,** Room MS1-084, Macdonald Stewart Building, 21111 Lakeshore Road, Ste. Anne de Bellevue, Quebec H9X 3V9. Telephone:514-398-7975, Fax: 514-398-7964

Journal of Clinical Microbiology 2010, 48:375-386

ABSTRACT

One of the major concerns in global public health and the dairy industry is the emergence of host-specific virulent *Staphylococcus aureus* strains. The high stability of the species genome renders detection of genetic microvariations difficult. Thus, approaches for rapid tracking specialized-lineages are urgently needed. We have used the clumping-factor A gene (*clfA*), in comparison to pulsed-field gel-electrophoresis (PFGE) and *spa*, for profiling of 87 bovine mastitis isolates from four regions in Canada. Twenty-five pulsotypes were obtained using PFGE with an index of discrimination of 0.91. These were assigned to six PFGE lineage-groups A-F and seven *spa*-types including two novel ones. Group A had 48.3% and D 43.7 % of isolates while only 8% were variable. MIC as determined by broth microdilution indicated that all isolates, except three, were susceptible to methicillin and to non beta-lactam antibiotics. All isolates had the *clfA* gene present and belonged to 20 *clfA*-repeat-types with an index of discrimination of 0.9. The dominant *clfA*-types X, Q, C, and Z formed 82% and 43% of PFGE groups A and D, respectively, and varied only within a narrow range of copy number of 46 to 52 copies, implying clonal-selection. The rest were uniquely variable and region-specific. Furthermore, the dominant groups contained subpopulations in different regions across Canada. Sequence information confirmed relatedness obtained by *clfA* repeat-copy-numbers and other methods and further revealed occurrence of "full-repeat" deletions and conserved host-specific codon-triplet position-biases at 18bp-units. Thus, concordant with PFGE and *spa*, *clfA* typing proved useful in revealing the clonal nature of mastitis-lineage and in rapid profiling subpopulations with comparable discriminatory power.

INTRODUCTION

Staphylococcus aureus have been a public health concern and a significant economic burden globally. In the dairy industry *S. aureus* bovine mastitis is one of the most important diseases with considerable economic losses (Seegers et al., 2003). The enterotoxigenic *S. aureus* strains, in dairy products, are major causes of foodborne disease in many countries (Asao et al., 2003; Villard et al., 2005). Although a limited number of dominant clones are responsible for the majority of infections (Kapur et al., 1995; Fitzgerald et al., 1997), there has been a rise in subtypes with elevated virulence and epidemicity in both hosts. A devastating example is the ongoing clonal expansion and diversification of a subset of community-associated methicillin-resistant *S. aureus* (CA-MRSA) isolates classified as USA300 (Kennedy et al., 2008; Klevens et al., 2007. Likewise, there is the emergence of a hyper-virulent strain, labelled ET3-1, a subtype of the common mastitis ET3 clone (Guinane et al., 2008). A potential risk therefore would be the parallel emergence of zoonotic infections (Juhasz- Kaszanyitzky et al., 2007; van Loo et al., 2007); particularly, the ET3-1 is highly prone to the acquisition of resistance determinants (Sung and Lindsay, 2007). The factors and mechanism(s) underlying the specificity and evolution of mastitis and CA-MRSA lineages have been quite elusive. Therefore, there is an urgent need for objective strain-subtyping tools that couple straightforward epidemiological investigations with pathogenesis and colonization potentials.

Multilocus sequence typing (MLST) grouped *S. aureus* in five main clonal complexes (CC) 8, CC30, CC5, CC22, and CC45 (Feil et al., 2003, 2004; Robinson and Enright, 2004; Enright et al., 2002) which were represented in three major and two minor clusters of amplified fragment length polymorphism (AFLP) (Melles et al., 2004). High-throughput genotyping of human and animal isolates has consistently shown a common background and virulence gene content, but mastitis-specialized *S. aureus* belonged to distinct clusters (van Leeuwen et al., 2005). Subsequent genomic analysis revealed 10 dominant lineages in each host (Lindsay et al., 2006; Sung et al, 2008). Similarly, mastitis-lineages were uniquely caused by ST151, ST771 and ST97, and unique genes conserved in all human isolates were variable or missing in animal lineages, including the

lineage-specific genes *fnbA, fnbB* and *coa*. It was suggested that a handful of genes account for host-adaptation (Sung et al., 2008). Hence, although stable core genome markers, such as MLST, are useful in identifying *S. aureus* common ancestors, the phenotypes of the same sequence type could be quite different. For instance, both strains MW2 and MSSA467 belong to ST1 but the former caused a fatal MRSA bacteraemia in USA, and the latter was isolated from an invasive MSSA osteomylitis in UK (Holden et al., 2004).

The species-specific variable genomic domain (*ori*C *environ*) is responsible for staphylococcal-speciation (Takeuchi et al., 2005; Dordet-Frisoni et al., 2007; Lindsay et al., 2006) and contain genes for adherence and invasion mediated by specialized structures called microbial-surface-component recognizing adhesive matrix molecules" (MSCRAMM) (Foster and Hook, 1998) such as *spa* and *clfA*. These MSCRAMMs mediate host-pathogen interactions (Projan and Novick, 1997; Foster and Hook, 1998) and therefore, could be used to identify strains with common host- and tissue-specific properties. For instance, *spa*-typing has been one of the most successful single genetic markers for typing *S. aureus* (Koreen et al. 2004). Recently, when *spa*-typing was examined as an alternative to PFGE typing of the Canadian MRSA epidemic clones, high concordance was found demonstrating the feasibility as a more rapid and simple but powerful alternative (Golding et al, 2008). However, the usefulness of *spa* typing for bovine isolates is questionable. SpA interaction in cows is not clear; it formed insoluble complexes with serum IgGs from Guinea pig and mouse only, but not from cow, goat, sheep, horse or chicken (2 Atkins et al., 20058) implying a different mechanism in hoofed animals.

In addition to *spa*, another major MSCRAMM class is the clumping factors, ClfB (*clfB*) and ClfA (*clfA*). The ClfB is expressed early in growth under increased oxygenation, and digested later by the stationary phase proteases (McAleese *et al*., 2001; Ni Eidhin et al., 1998). This questions its role in deep-seated oxygen-limited microenvironments, such as intra-mammary (IMI) infections- a condition known to induce emergence of persistent variants as the small colony types, where significant ClfA activity has been shown by IMI challenge of mice immunized with anti-ClfA antibody (Tuchscherr et al., 2008). On the other hand, several properties in the latter make it an

ideal tool for specialized clones. ClfA is constitutive, independent of the Agr-system, abundant in deep-infections (Wolz et al., 2002; Josefsson et al., 2008), and mediates virulence even in absence of fibrinogen (Palmqvist et al., 2004). Further, the number of repeats in the *clfA* affected the adherence and clumping titers of the cocci *in vitro* (Risley et al., 2007; Hartford et al., 1997). Moreover, ClfA and not fibronectin binding proteins was responsible for the intracellular counts of the cocci (Ahmed et al., 2001), and for the bacterial load and dissemination leading to abscess formation (Cheng et al., 2009; Stranger-Jones et al., 2006; Nanra et al., 2009). Thus, *clfA* is a useful strain-specific marker that is reasonable for screening specialized lineages. We have recently evaluated the usefulness and stability of this marker using isolates and strains from human infections and bovine mastitis (Said et al., 2009a). In addition, we have further provided new perspectives on the organ- and host-specificity of *clfA* from temporal and geographically independent clinical isolates of different human organs and bovine mastitis (Said et al., 2009b).

The last study on genetic structure and antimicrobial susceptibility of *S. aureus* recovered from cows in Canada was carried out on isolates from 1999-2000 in Ontario and Quebec (Sabour et al., 2004). Of these isolates tested for their resistance to antimicrobials 24.5% were resistant to at least one antimicrobial with resistance to penicillin being the most common (9.9%), followed by sulfadimethoxine (7.5%). Isolates collected in the province of Ontario exhibited the highest proportion of resistant isolates (30.2%). In Quebec, only resistance to penicillin, tetracycline, and sulfadimethoxine was found. Resistance to the penicillin-novobiocin combination, cephalothin, or ceftiofur was not found in any isolate (46). Despite the *in vitro* susceptibility, it is still difficult to eradicate the pathogen due to invasiveness and the development of intracellular dormant variants.

Many recent comparative genomic analysis have consistently reached the same conclusion, that being *S. aureus* from humans and animals have the same genetic background and gene content (van Leeuwen et al., 2005; Lindsay et al., 2006) and that the differences in virulence in the seemingly identical emerging clones are due to a few subtle changes rather than to large-scale acquisition of virulence factor genes (Highlander et al., 2007; Sung et al., 2008; Kennedy et al, 2008). The rapidly changing epidemiology and

evolution necessitates constant strain-specific regional profiling. Thus, the objective of this study was to exploit *clfA* for strain diversity in Western, Central, and Eastern Canadian provinces, in comparison with PFGE and *spa* as well as the antimicrobial susceptibility of those isolates.

MATERIALS AND METHODS

Bacterial strains

A total of 87 *S. aureus* isolates (Table 4.1; CO15 and CE22 were examined only by *clfA*) from clinical mastitis in 78 cows that were housed in 24 different farms from different regions across Canada were studied. The geographic distribution of the isolates, which were collected from January 2007 to April 2008, was as follows: 6 farms located in western Canada, 5 farms in Ontario, 6 farms in Quebec, and 7 farms in eastern Canadian provinces. Thus, except for nine cows that had two independent isolations at two different timings, all isolates were from different cows. These were obtained from the Canadian Bovine Mastitis Research Network (CBMRN) database, kindly provided by Dr. Grant Tomita. They were retested by the STAPH LATEX KIT (Pro-Lab Diagnostics Canada, Ontario, Canada) and further confirmed by the identification of the *nuc* gene. In addition, published sequences of six strains RF122, Newman, COL, MSSA476, and the USA300 strains *Staphylococcus aureus* subsp. aureus USA300 PR3757 (abbreviate as FPR) and *Staphylococcus aureus* subsp. aureus USA300 TCH1516 (abbreviated as TCH) were used as references for sequence analysis. Glycerol stocks of isolates in trypticase soy broth (TSB) were stored at -80°C.

Preparation of genomic DNA

PFGE was carried out using *Sma*I at the Laboratoire de santé publique du Québec (LSPQ) according to the Canadian Standardized Protocol for PFGE of *S. aureus* (Mulvey et al., 2001) SOP number ARNI-PR-001, National Microbiology Laboratory (NML), Winnipeg, Canada.

Macrorestriction fingerprint patterns were analyzed by Bionumeric software v5.0 (Applied Maths , Austin, Tx, U.S.A) and dendrograms were created by using the Dice

similarity coefficient and the unweighted pair group method with arithmetic means. The band position tolerance of 1.0% was used, and the cluster cutoff was set at an 80% similarity level. Identical restriction patterns were assigned to the same type, whereas types that differed from the common type by 1-6 band differences or less were assigned to the same lineage group, according to Tenover et al., criteria (1995). Lineage groups were named alphabetically, and types within those groups were assigned numerically.

Antimicrobial susceptibility testing

Susceptibility to nine antibiotics was examined at the (LSPQ) in TC microwell 96-well round-bottom U-Plates (Fisher Scientific, Ottawa, ON, Canada). These were: Oxacillin, cephalotin, tetracycline, erythromycin, penicillin, trimethoprim-sulfamethoxazole, penicillin-novobiocin. Wells in plates were inoculated with Mueller-Hinton broth cultures of the test isolates and as recommended by the manufacturer in accordance with the Clinical Laboratory Standards Institute (CLSI 2008) guide M31-A3 for veterinary isolates. After incubation, plates were read by eye for direct observation of growth or no growth. Isolates were scored as antibiotic sensitive, intermediate, or resistant based on growth or no growth at the appropriate breakpoint MIC for a given antibiotic based on CLSI guide; the breakpoints used are given in the results. Reference strain *S. aureus* ATCC29213 served as the assay control.

***clfA* typing**

PCR amplification of *clfA*R domain typing was carried out as described previously (Said et al., 2009a; 2009b). Differences in R-domain copy numbers were obtained from PCR product sizes on gels using information on *S. aureus* COL. Repeat-types (RTs) were assigned based on difference in the 18bp copy numbers; a difference by a single (18bp) copy was considered a RT. It is widely accepted that isolates within one PFGE pattern are genetically identical, based on this we first determined dominant PFGE pattern groups, then for the assumption that variant isolates have evolved from the dominant one, we had identified the variant RTs for *clfA* within these groups.

clfA-sequencing, sequence analysis, and phylogenetic relationships

The copy numbers were confirmed by sequencing the *clfA*-R domain PCR amplicons at McGill University and Quebec Genome Innovation center. *clfA* nucleotide sequence analysis and phylogenetic relationships were carried out according to Said et al., (2009a). We have used the whole-repeated region sequences along with conserved 3' and 5' regions for alignments and phylogenetic relationships as a measure of "relatedness". Global multiple sequence alignments of the 26 pulsotypes were carried out using Geneious Bioinformatics package v4.6 with most common transversions highlighted. Sequences were further checked visually. Agreement, disagreement, transitions, and transversion were also checked by the program using a 100% nucleotide identity against a common consensus. Phylogenetic grouping was carried out with the help of the Neighbor-Joining method built in the Geneious bioinformatics package v4.6. The phylograms were based on the most likely groupings that reflect genetic relationships in circular trees.

clfA discriminatory power

Determination of RTs was based on differences in the number of the 18-bp copies of tandem repeats contained in the R-domain of the *clfA* gene. The Discriminatory Power Calculator (http://biophp.org/stats/discriminatory_power/demo.php) was used to calculate the index of discrimination (ID); where a value of 1 indicates the ability to differentiate each isolate, and a value of 0 indicates that all isolates are identical. The numerical index of discriminatory power was used to give numeric estimates for strain differentiation and the values (defined as the average probability that the typing system will assign a different type to two unrelated strains randomly sampled in the microbial population of a given taxon) estimated according to Hunter, et al. (1990). The different types of organization of a repeat region termed repeat profiles, was not used in this study as we attempt to correlate relationships based on repeat "unit" differences as well as sequence information with an ultimate goal of introduction of a rapid primary screening tool with repeats without sequencing as applied to tandem-repeat typing systems (van Belkum et al., 2007).

Real-time PCR assay

The RT-PCR assay for *mecA, nuc,* and *lukPV* was performed on the PFGE genotypes at the NML using primers, protocols and instruments, according to McDonald et al., (2005).

spa sequencing

PCR amplification of the *spa* repeat region was performed as previously described (Harmsen et al., 2003). Amplicons were sequenced in house by the DNA Core Facility at the NML. The DNA sequences of the *spa* repeat region in both directions were imported as ABI or SCF files and analyzed using the *spa* typing program provided with BioNumerics v5.0 (Applied Maths). DNA sequences were compared through use of the *spa* typing websites http://tools.egenomics.com/public/login.aspx and http://www.*spa*server.ridom.de, the latter of which was developed by Ridom GmbH and curated by SeqNet.org (http://www.SeqNet.org/).

RESULTS

Pulsed-field gel electrophoresis

SmaI macrorestriction fragments of 87 isolates produced 25 patterns (Figure 4.1) with an ID of 0.91. Based on the cluster cutoff set at 80% similarity level, these were assigned to six lineage groups designated A to F. Lineage group A was the most common to which 42 (48.3%) of the 87 isolates belonged. Subgroup A1 comprised about half (20 isolates) of the isolates that were, with the exception of six, from the Quebec region. The other half of isolates (22 isolates) were all, except four, from Eastern Canada and formed the distinct region-specific clonal lines namely A3 (8 isolates), A6, A7, A8, and A9. Sublineage A4 included four isolates (two from East and two from West), A5 was restricted to Ontario. Of the remaining 45 isolates, 38 (43.7%) made up the second largest lineage, labelled group D, to which all Western isolates belonged, except four. Sublineage D1 was the largest with two isolates from each region but was dominated by Western isolates, followed by D7 which was unique only to West. However, sublineage D10 was

unique to the Eastern region. The reminder of seven isolates belonged to groups B with one isolate from East and another from West, C which contained all three isolates from Ontario, in addition to unique lineages E and F from East and Quebec, respectively. Thus, each region had certain sublineage patterns either with high frequency, such as PFGE pattern 19 and 18 in the West, pattern 3 in East, and pattern 1 in Quebec or with moderate frequency such as pattern 20 in Ontario. There were also some types found in all regions such PFGE pattern 1 and 19. Genetic heterogeneity within farms was also found. As shown in Table 4.1, the majority of the isolates showed alpha-beta double-hemolysis with larger beta zones on BA.

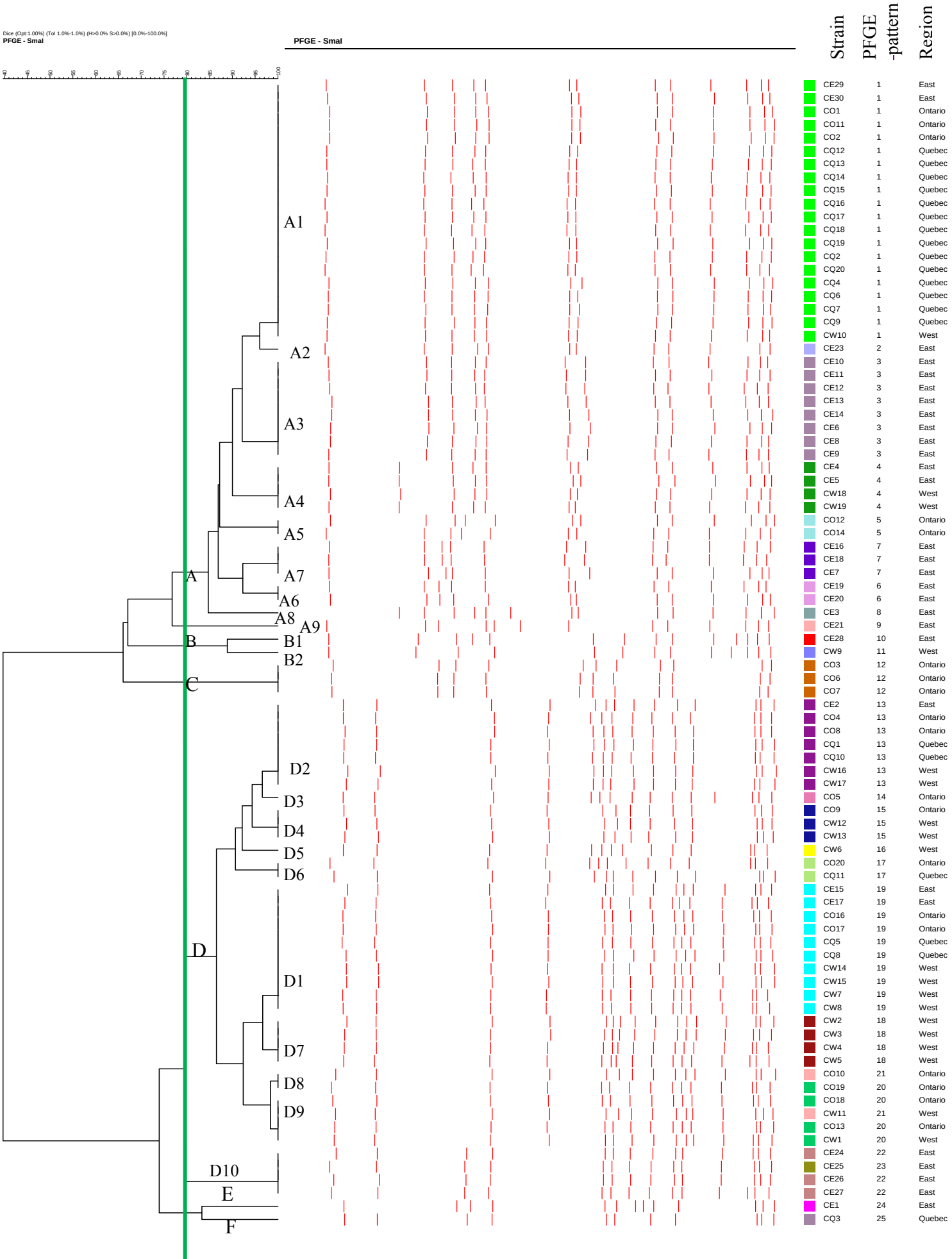

Dice (Opt:1.00%) (Tol 1.0%-1.0%) (H>0.0% S>0.0%) [0.0%-100.0%]
PFGE - SmaI
PFGE - SmaI
Strain
PFGE pattern
Region
A1
A2
A3
A4
A5
A7
A6
A8
A9
B1
B2
D2
D3
D4
D5
D6
D1
D7
D8
D9
D10
A
B
C
D
E
F
CE29 1 East
CE30 1 East
CO1 1 Ontario
CO11 1 Ontario
CO2 1 Ontario
CQ12 1 Quebec
CQ13 1 Quebec
CQ14 1 Quebec
CQ15 1 Quebec
CQ16 1 Quebec
CQ17 1 Quebec
CQ18 1 Quebec
CQ19 1 Quebec
CQ2 1 Quebec
CQ20 1 Quebec
CQ4 1 Quebec
CQ6 1 Quebec
CQ7 1 Quebec
CQ9 1 Quebec
CW10 1 West
CE23 2 East
CE10 3 East
CE11 3 East
CE12 3 East
CE13 3 East
CE14 3 East
CE6 3 East
CE8 3 East
CE9 3 East
CE4 4 East
CE5 4 East
CW18 4 West
CW19 4 West
CO12 5 Ontario
CO14 5 Ontario
CE16 7 East
CE18 7 East
CE7 7 East
CE19 6 East
CE20 6 East
CE3 8 East
CE21 9 East
CE28 10 East
CW9 11 West
CO3 12 Ontario
CO6 12 Ontario
CO7 12 Ontario
CE2 13 East
CO4 13 Ontario
CO8 13 Ontario
CQ1 13 Quebec
CQ10 13 Quebec
CW16 13 West
CW17 13 West
CO5 14 Ontario
CO9 15 Ontario
CW12 15 West
CW13 15 West
CW6 16 West
CO20 17 Ontario
CQ11 17 Quebec
CE15 19 East
CE17 19 East
CO16 19 Ontario
CO17 19 Ontario
CQ5 19 Quebec
CQ8 19 Quebec
CW14 19 West
CW15 19 West
CW7 19 West
CW8 19 West
CW2 18 West
CW3 18 West
CW4 18 West
CW5 18 West
CO10 21 Ontario
CO19 20 Ontario
CO18 20 Ontario
CW11 21 West
CO13 20 Ontario
CW1 20 West
CE24 22 East
CE25 23 East
CE26 22 East
CE27 22 East
CE1 24 East
CQ3 25 Quebec

Figure 4.1. Pulsed-field gel electrophoresis of *SmaI*-digested total DNA from *S. aureus* isolates. The dendrogram shows estimates of genetic relatedness of the PFGE types. The 25 PFGE patterns were dividable into six major lineage groups, designated by the letters A through F, and these were further subdivided into sublineages indicated by numerals associated to group letters. The colored boxes highlight the PFGE pattern types.

***clfA*-typing**

All of the 87 isolates were also analyzed for the *clfA* repeat region, and all had the gene present. *clfA* revealed 20 different repeat types (RTs) designated A to R, X and Z (Table 4.1) with an ID of 0.9. There was concordance between *clfA* and PFGE grouping of major clusters and independent lineage assignments. However, *clfA* further distinguished new types within PFGE groups in many occasions, except for a few instances where PFGE was superior, as shown in Table 4.3. An important observation is that 82% of isolates in PFGE lineage group A belonged to the dominant RTs X, Q, C, and Z, further subdividing the group A into 10 different RTs (Table 4.1). Of the 42 PFGE group A isolates, 14 were RT of X and all except two were PFGE pattern 1. Nine were of type Q and all, except 2 were of pattern 3, seven were of type C, and five were type Z. The latter two RTs had mixed PFGE patterns, and the rest of repeat types F, R, G, H, I, and N were under represented. As indicated in Table 4.1, RT of X were from independent cows at four different farms except for CQ14/CQ15 and CE20/CE21 which were (each pair) from the same cow but each cow had two independent isolations at two different times, and from different quarters or teats. The haemolytic pattern of the first pair of isolates was 4/2 as shown in Table 4.1. Repeat-type Q, C, and Z were also from different cows located at different regions except CQ9/CQ10 which were two independent isolations from a cow, and CE4/CE5 were two isolations on the same date from different quarters of a cow. The same major RTs, X, Q, C, and Z were also present, but in lower frequency than PFGE group A, in the PFGE group D making 43% (out of 38). The rest of the PFGE group D were mostly unique RTs. The remaining seven PFGE groups were also represented by unique RTs. Major RTs in Group D were characterized by representations of isolates from wider geographic regions except for the new RT of A

restricted to a farm in Eastern region. Each region had one or two minor RTs of 2-3 isolates from different cows in the same herd, and showed different PFGE patterns.

Table 4.1. Comparative analysis of *clfA* types with PFGE and *spa*-typing.

'Isolates' column shows isolates used. First letter 'C' in isolate code stands for Clinical mastitis/Canada, and the second letter is the first letter of each province Q for Quebec, O for Ontario, W for West, and E for Eastern provinces. The number before "period" is that isolate number, and after "period" is the farm number. 'Hemolysis' is given letter codes as follows; 1 = beta, 2 = weak beta, 3 = non-hemolytic, 4 = double hemolytic zones (clear and incomplete hemolysis rings). 'Copy number' column shows *clfA* repeat copies for that isolate as obtained by sequencing PCR products. Columns showing 'RT' and 'PFGE lineage' indicate Repeat-Types and PFGE lineage groups, respectively, designated by letters. The column showing 'PFGE patterns' indicate PFGE pattern numbers. Columns 'Ridom- and Kreiswirt- *spa'*, show standard designations for *spa* types obtained.

Isolates	Hemolysis	Copy number	RT	PFGE Lineage	PFGE pattern	Ridom *spa*	Kreiswirt *spa*	Isolates	Hemolysis	Copy number	RT	PFGE Lineage	PFGE pattern	Ridom	Kreiswirth	Isolates	Hemolysis	Copy number	RT	PFGE Lineage	PFGE pattern	Ridom	Kreiswirth
CQ2.2	4	52	X	A	1	t267	*spa*105	CO12.4	4	50	Z	A	5			CE27.7	4	54	A	D	22		
CQ4.3	4	52	X	A	1			CW18.	4	50	Z	A	4			CW13.4	1	39	K	D	15	t529	*spa*102
CQ9.5	4	52	X	A	1			CO14.4	4	50	Z	A	5	t529	*spa*102	CW14.4	4	39	K	D	19	t529	*spa*102
CQ12.5	4	52	X	A	1			CE29.7	4	50	Z	A	1			CW15.4	4	39	K	D	19		
CQ13.5	4	52	X	A	1			CE30.5	4	50	Z	A	1			CE15.5	4	49	D	D	19		
CQ14.5	4	52	X	A	1			CW19.	4	47	F	A	4			CE17.5	4	49	D	D	19		
CQ15.5	2	52	X	A	1			CE14.5	4	47	F	A	3			CW11.3	4	36	M	D	21		
CQ16.5	4	52	X	A	1			CO9.3	4	21	R	A	15			CW12.3	4	36	M	D	15		
CQ17.5	4	52	X	A	1			CO11.3	4	21	R	A	1			CO8.3	4	21	R	D	13		
CQ18.5	4	52	X	A	1			CW10.	4	21	R	A	1			CO10.3	4	21	R	D	21		
CQ19.5	4	52	X	A	1			CO2.1	4	45	G	A	1			CO13.4	4	50	Z	D	20		
CQ20.5	1	52	X	A	1			CE23.6	4	43	H	A	2	t244	new	CE2.2	4	50	Z	D	13		
CE20.6	4	52	X	A	6	t3380	*spa*106	CE7.2	4	41	I	A	7			CO4.1	4	53	B	D	13		
CE21.6	4	52	X	A	9	t3380	*spa*106	CO1.1	1	34	N	A	1			CW16.5	4	47	F	D	13		
CE22.6	4	52	X	ND	ND			CQ1.1	1	46	Q	D	13	t529	*spa*102	CO16.5	4	45	G	D	19		
CE6.4	4	46	Q	A	3	t2445	new	CW4.2	4	46	Q	D	18	t605	*spa*102	CO17.5	4	40	J	D	19		
CE8.4	4	46	Q	A	3			CW6.2	4	46	Q	D	16	t605	*spa*102	CO18.5	4	40	J	D	20	t529	*spa*102
CE9.4	4	46	Q	A	3			CW7.2	4	46	Q	D	19	t605	*spa*102	CW8.2	4	37	L	D	19		
CE10[4]	4	46	Q	A	3			CQ11.5	4	46	Q	D	17	t529	*spa*102	CO19.5	4	29	O	D	20		
CE11.4	4	46	Q	A	3			CW2.1	4	51		D	18			CO6.2	4	46	Q	C	12		
CE12.4	4	46	Q	A	3			CW5.2	4	51	C	D	18			CO7.2	4	46	Q	C	12		
CE13.4	4	46	Q	A	3			CQ5.4	4	51	C	D	19			CO3.1	4	53	B	C	12	t529	*spa*102
CE16.5	4	46	Q	A	7			CQ8.4	1	51	C	D	19			CW9.3	4	37	L	B	11	t529	*spa*102
CE18.5	4	46	Q	A	7	t2445	new	CQ10.5	4	51	C	D	13			CE28.7	4	27	P	B	10	Unk	New
CQ6.4	2	51	C	A	1			CO5.1	4	52	X	D	14	t359	*spa*92	CE1.1	4	53	E	E	24	t529	*spa*102
CQ7.4	1	51	C	A	1			CW1.1	4	52		D	20	t521	*spa*88	CQ3.2	4	48	F	F	25	t529	*spa*102
CO15.4	1	51	C	ND	N			CW3.1	4	52	X	D	18										
CE4.3	4	51	C	A	4			CW17.	4	52	X	D	13										
CE19.6	4	51	C	A	6			CE24.7	4	54	A	D	22	t529	*spa*102								
CE3.3	4	51	C	A	8	t267	*spa*105	CE25.7	4	54	A	D	23	t529	*spa*102								
CE5.3	4	51	C	A	4	t267	*spa*105	CE26.7	4	54	A	D	22										

Table 4.2. *spa*-types, its repeat successions, and RealTime assay of *mecA*, *nuc*, and *pvl* genes from the isolates of 25 PFGE patterns. The variable mosaic of recombinogenic repeat types are shaded

Code	*mecA*	*nuc*	*pvl*	Ridom	RepeatSuccession	Kreiswirth	Repeat	Seen in ND	PFGE	RT
CE25	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	A
CE24	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	A
CW13	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	K
CW14	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	K
CO18	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	J
CQ1	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	Q
CQ11	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	D	Q
CW4[2]	-ve	+ve	-ve	t605	r07r23	Spa102	UJ	t605 no	D	Q
CW6	-ve	+ve	-ve	t605	r07r23	Spa102	UJ	t605 no	D	Q
CW7	-ve	+ve	-ve	t605	r07r23	Spa102	UJ	t605 no	D	Q
CO5	-ve	+ve	-ve	t359	r07r23r12r21r17r34r34r33r34	Spa92	UJGFMBBPB	t359 ND with ST97	D	X
CW1	-ve	+ve	-ve	t521	r07r23r12r21r17r34r34r34r34r33r34	Spa88	UJGFMBBBBPB	t521,ND with ST97	D	X
CQ2	-ve	+ve	-ve	t267	r07r23r12r21r17r34r34r34r33r34	Spa105	UJGFMBBBPB	t267ND with ST97	A	X
CE20	-ve	+ve	-ve	t3380	r07r23r12r21r17r34r34	Spa106	UJGFMBB	t3380 no	A	X
CE21	-ve	+ve	-ve	t3380	r07r23r12r21r17r34r34	Spa106	UJGFMBB	t3380 no	A	X
CE6	-ve	+ve	-ve	t2445	r07r16	new	UK	t2445 no	A	Q
CE18	-ve	+ve	-ve	t2445	r07r16	new	UK	t2445 no	A	Q
CE3	-ve	+ve	-ve	t267	r07r23r12r21r17r34r34r34r33r34	Spa105	UJGFMBBBPB	t267ND with ST97	A	C
CE5	-ve	+ve	-ve	t267	r07r23r12r21r17r34r34r34r33r34	Spa105	UJGFMBBBPB	t267ND with ST97	A	C
CO14	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	A	Z
CE23	-ve	+ve	-ve	t2445	r07r16	new	UK	t2445 no	A	H
CW9	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	B	L
CE28	-ve	+ve	-ve	Unkno	r07r23r12r12r21r17r34r34r34r34r33r34	new	UJGGFMBBBBPB	Submitted to	B	P
CO3	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	C	B
CE1	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	E	E
CQ3	-ve	+ve	-ve	t529	r04r34	Spa102	ZB	t529 no	F	F

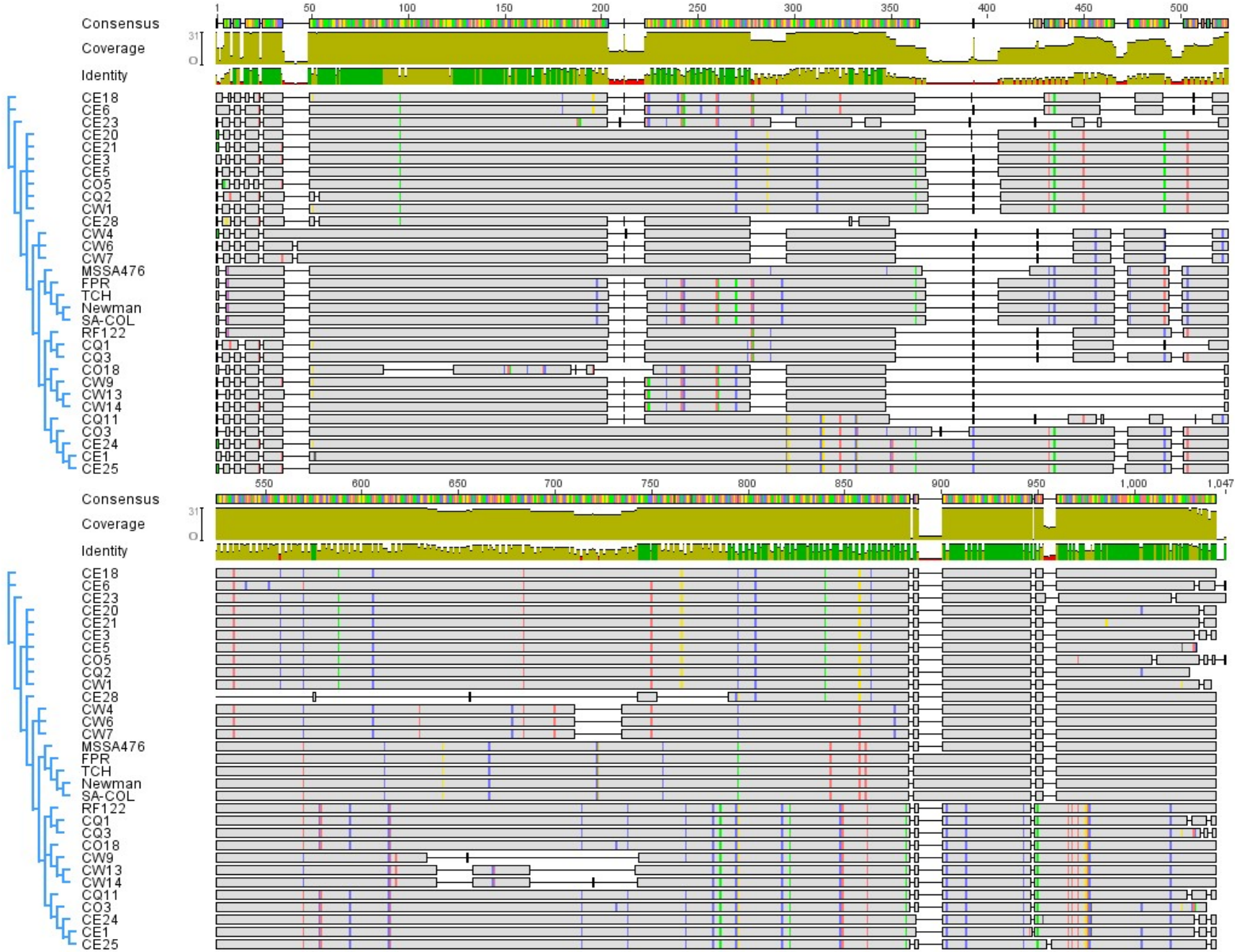
Table 4.3. PFGE lineage groups and the corresponding RTs. Numbers in bracket indicate number of isolates for that type.

PFGE group	PFGE pattern	Corresponding *clfA* types	New/recombinogenic *clfA* types
A1	1 (20)	X(12)	C(2), Z(2), R(2), G(1), N(1)
A3	3 (8)	Q(7)	F(1)
A4	4 (4)	C (2)	Z(1), F(1)
A5	5 (2)	Z(2)	
A6	7 (3)		Q(2), I(1)
A7	6 (2)		X(1), C(1)
C	12(3)		Q(2) B(1)
D2	13(7)		Q(1), C(1), X(1), R(1), Z(1), B(1), F(1)
D4	15(3)		R(1), K(1), M(1)
D1	19(10)	K(2), D(2), J(1),	Q(1), C(2) L(1) G(1),
D7	18(4)		Q(1), X(1), C(2)
D10	22/23(4)	A(4)	
D8, D9	20/21(6)	O(1)/M(1), R(1)	X(1), Z(1), J(1)
A2	2(1)	H	
A8	8(1)	C(1)	
A9	9(1)	X(1)	
B1, B2	10(1)/11(1)	L(1), P(1)	
D3	14(1)		X(1)
E	24(1)	E	
F	25(1)	F	
D5	16(1)		Q(1)
D6	17(2)		Q(1), ND(1)

Because of the widely established notion that *S. aureus* is clonal both in core- as well as in polymorphic-genes (Feil et al., 2003; Kuhn et al., 2006), diversifications are assumed to occur more by point mutation rather than by recombination. To confirm copy numbers, to establish evolutionary relationships, and to examine which type of mutations (point mutation or recombination) would be the most likely driving force for repeat variations, we analyzed the whole R domain sequences with 3' and 5' conserved regions of the 25 PFGE patterns obtained. The copy numbers of *clfA* were confirmed by sequencing. Repeat-types with identical or similar copies had identical or very similar nucleotide alignment patterns (Figure 4.2A). These groupings were highly consistent with groupings made by both other methods. Many isolates with either identical or similar

repeat copies which belonged to the same or similar PFGE lineage groups, showed the same nucleotide alignment patterns along the whole R domain. For instance, the PFGE lineage group A isolates CQ2, CE3, CE21, CE20, CO5, and CE5, all except the last two which were of similar RT of C, were of type X and had same sequence type (Figure 4.2A). Similarly, the following two groups of isolates had two distinct patterns, CW4, CW6, CW7 that were RT of Q, (PFGE group D) and CW9, CW13, CW14 were RT of L for the first and K for the last two (lineage groups B and D, respectively). Finally, the USA300 subsets FPR and TCH were closely related to COL and Newman and the MSSA isolate MSSA476 showed a relative evolutionary distant within the group (Figures 4.2A and 4.3). All had 51 copies of repeats except the MSSA476 which had 49 copies.

Figure 4.2A also shows that the overall patterns were highly conserved within related groups except in the middle of the R domain sequence where extensive full "repeat-unit" deletions were very common, and that related groups had very closely related deletion patterns. In addition, all of the sequences analyzed (human or animal) had the same 18-bp repeat-unit structures that makeup the 6 codon-motifs with a highly conserved positions suggesting strong codon usage preference at those sites. Furthermore there were also a few repeat-units with varying lengths. Additionally, human and animal isolates uniquely selected different nucleotide codon triplets for the same amino acid at particular positions, one of these regions is shown in Figure 2B. In this figure, for animal strains the repeat unit number 1 [**TCA** GAC **TCA** GAC **AGC** GAC] shows the less conserved form of ser codon in the 1st, 3rd, and 5th positions (**bold**). In subsequent units 2, 3, 4, and 5 the more conserved form of ser codon (AGT) is always in the 5th position, while the GAC form of aspirate in the 2nd, 4th, and 6th positions. For human strains [**TCC** GAC **TCC** GAC **AGT** GAC] there is codon bias towards the more conserved primordial forms of ser (**bold**).This is supported by an additional U-rich bias in asparate codon (GAT).


Consensus
Coverage
Identity
CE18
CE6
CE23
CE20
CE21
CE3
CE5
CO5
CQ2
CW1
CE28
CW4
CW6
CW7
MSSA476
FPR
TCH
Newman
SA-COL
RF122
CQ1
CQ3
CO18
CW9
CW13
CW14
CQ11
CO3
CE24
CE1
CE25

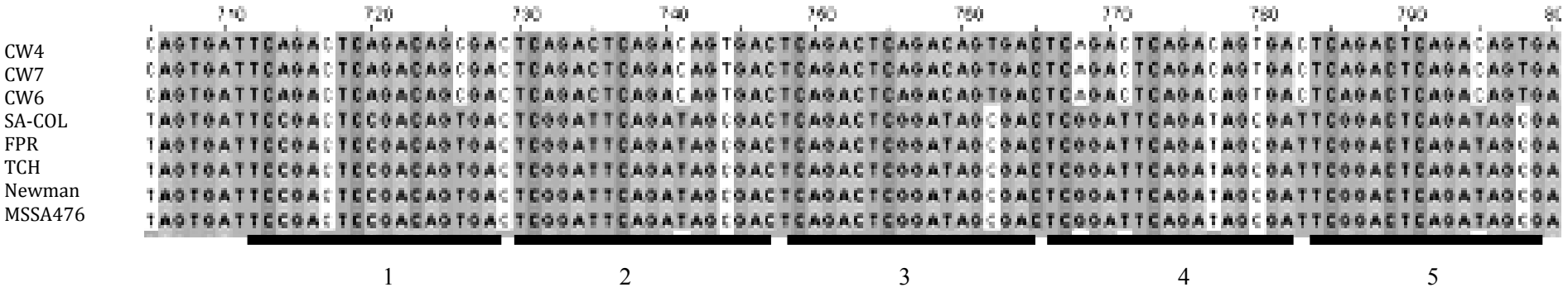


Figure 4.2 A. Whole *clfA*-R domain global multiple nucleotide sequence alignment patterns of isolates of the 25 PFGE patterns (Geneious bioinformatics package v4.6) with Cost-Matrix of 65% similarity (5.0/4.0) and free end gaps. A combined guide tree based on Neighbor-Joining (TN distance model) was used. Common transversions were highlighted by base colors as follows: Green = T, Blue = C, Red = A, and Yellow = G. Identity and sequence coverage are indicated by color, and the most common consensus sequence was selected. Standard strain abbreviations were used; *Staphylococcus aureus* subsp. aureus USA300 PR3757 (FPR), *Staphylococcus aureus* subsp. aureus USA300 TCH1516 (TCH), *Staphylococcus aureus* subsp. COL, (COL). Figure 2B. Local multiple sequence alignment of selected human and animal isolates showing only one region selection of amino acid codon triplets by isolates from the two hosts. In repeat units 1 to5, the nucleotide codon triplets in animal/human strains, respectively are shown in the following patterns: repeat unit 1 ser AGC/AGT, repeat unit unit 2 GAC/GAT, GAC/GAT, AGT/AGC, repeat unit 3, GAC/GAT, AGT/AGC, repeat unit 4 GAC/GAT, GAC/GAT, AGT/AGC, GAC/GAT, and in repeat unit 5 GAC/GAT, AGT/AGC.

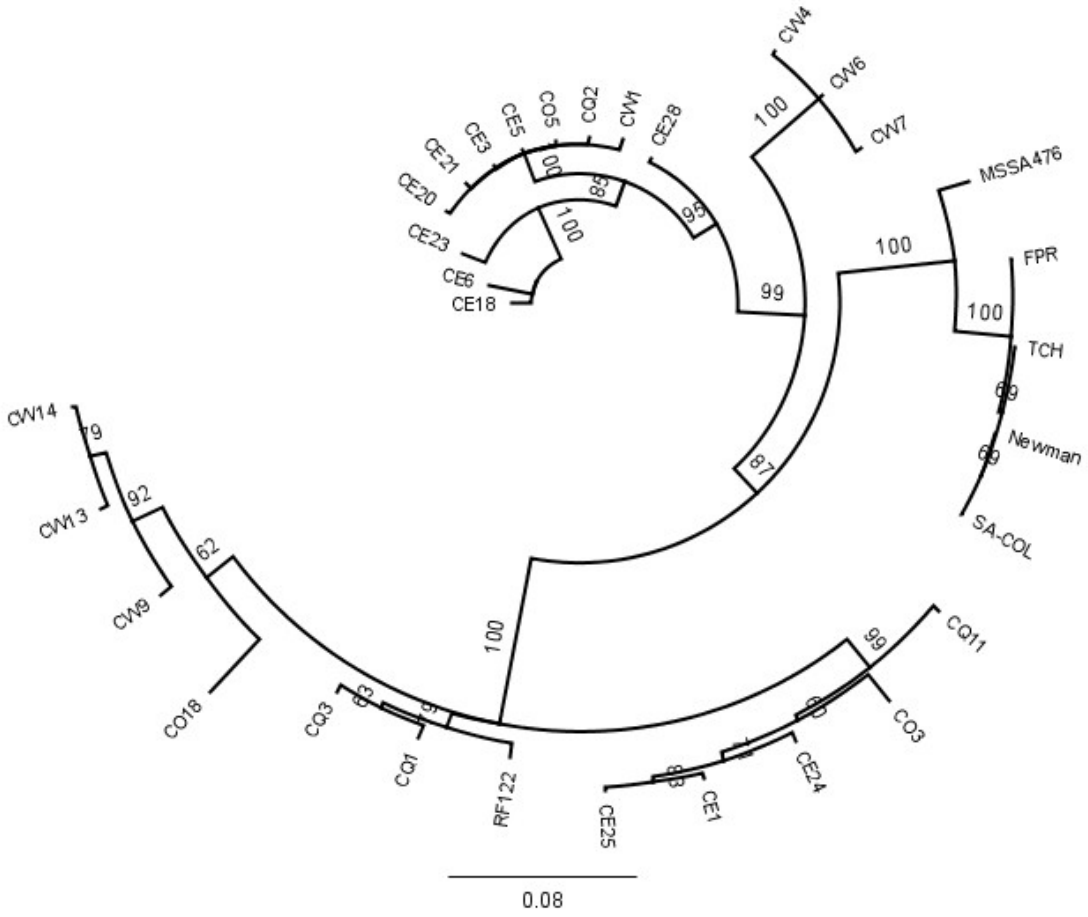


Figure 4.3. The phylogenetic relationships of isolates. Phylogram of the *clfA*-R domain sequences with conserved 3' and 5' regions based on the Neighbor-Joining program in Geneious bioinformatics package v4.6. Distances were estimated by the Tamura–Nei model using bootstrap as resampling method with support threshold of 50%.

spa-typing

Combination of PFGE and *spa* typing has been proved useful and in many typing instances these methods, particularly *spa* typing, also identifies the most common ancestor lineages. Thus, in this study the 26 isolates representing the 25 PFGE patterns were further *spa*-typed to examine the distribution of major lineage lines. These belonged to seven *spa*-types as identified by Ridom/ Kreiswirth systems; four types identified in PFGE lineage group A (t267/*spa*105, t3380/*spa*106, t529/*spa*102, and t2445/new (Kreiswirth repeat succession UK)),

and three other *spa*-types belonged to PFGE lineage group D (t529/*spa*102, t521/*spa*88, and t359/*spa*92). A novel type was found in lineage group B2 by Ridom (repeat succession r07r23r12r12r21r17r34r34r34r34r33r34) and Kreiswirth (repeat succession UJGGFMBBBBPB). In addition, Ridom types t605 and t529 were same by Kreiswirth (*spa*102).Thus, two new *spa*-types were identified as shown in Tables, 4.1 and 4.2. *spa* type t529/*spa*102 was predominant in group D, however, lineage groups B1, C, E, and F were indistinguishable by *spa*-typing as they all were type t529/*spa*102. In PFGE lineage group A, major *clfA* types were mostly associated with only one *spa* type. Similarly, in the more variable PFGE lineage group D, potentially recombinogenic *clfA* types Q and X, had more than one *spa*-type each. These groups also had the most variable PFGE pattern types.

Antimicrobial susceptibility test

All of the 87 *S. aureus* isolates used in this study were tested for their resistance to nine antimicrobials (Table 4.4). Three isolates, CO12, CO13, and CO14, were resistant to penicillin (all from Ontario) and one isolate, CE18, was resistant to tetracycline (Eastern Canada region). Erythromycin intermediate resistance was common in Eastern regions. In addition, the multiplex RT-PCR assay for *mecA,* and *lukPV* revealed that the *lukPV* and *mecA* genes were absent in all of the tested isolates (Table 4.2). Furthermore, screening for hemolysis revealed that the overwhelming majority produced mainly both types of hemolysis, clear and incomplete, on BA. This was irrespective of region or subregion as shown in Table 4.1.

Table 4.4. Antimicrobial breakpoints and number of isolates showing sensitive, intermediate, or resistance patterns in 87 mastitis *S. aureus* collected from dairy herds across Canada.

Antimicrobial Agent	Breakpoints* μg/ml			No of isolates sensitive (S), intermediate (I), or resistant (R) to the antibiotics											
	S	I	R	East			West			Quebec			Ontario		
				S	I	R	S	I	R	S	I	R	S	I	R
Oxacillin	≤2	-	≥4	29	0	0	19	-	0	20	-	0	19	0	0
Cephalotin	≤8	16	≥32	29	0	0	19	0	0	20	-	0	19	0	0
Tetracycline	≤4	8	≥16	28	0	1	19	0	0	20	-	0	19	0	0
Erythromycin	≤0.5	1-4	≥8	16	13	0	19	0	0	16	4	0	14	5	0
Penicillin	≥0.12	-	≥0.25	29	-	0	19	-	0	20	-	0	16	-	3
Trimethorpim-	≤2/38	-	≥4/76	29	-	0	19	-	0	20	-	0	19	-	0
Penicillin-novobiocin	≤1/2	2/4	≥4/8	29	0	0	19	-	0	20	-	0	19	-	0

*A (-) is shown where breakpoints were determined as either sensitive or resistant only.

DISCUSSION

In the light of the widely established clonal genetic background and gene content of *S. aureus* of different origins, the evolution of strains with higher epidemicity and virulence is quite elusive. Combinations of typing approaches that couple epidemiological data and strain phenotype might serve as a paradigm strategy, in both basic and applied research, to study the basis for specialization and evolution. One such method would be coding tandem repeats, particularly, in places where *S. aureus* infection is endemic. Recently, we have provided new perspectives on the use of *clfA* locus in organ- and host-specific grouping of clinical isolates from different host sites (Said et al., 2009b). In the present study, we have used the *clfA* typing to show for the first time the distinct subpopulations of repeat-profiles of mastitis-specific *Staphylococcus aureus* lineage in different regions in Canada in concordance with PFGE and *spa* grouping. The

general agreement among three typing methods suggests that the *clfA* typing is a useful strain screening tool in both basic and applied research.

In this study, concordance between PFGE lineage grouping, *spa*, and *clfA* types was obtained. The simplicity, reproducibility, and high discriminatory power of *clfA* typing make it a valuable tool for screening clonal groups of mastitis-specialized lineage before application of major methods. The dominance of the major RTs within PFGE lineage group A and the lower frequency prevalence, of the same RTs, in the completely different PFGE lineage group D and other groups (Table 4.1) is potentially an interesting indication of the recombinogenic nature of the locus in a clonal background, consistent with the finding of Koreen et al., (2005) on the *clfB* locus. Despite the fact that in clonal diversification, point mutations within MLST loci give rise to new alleles at least 15-fold more frequently than does recombination (Feil et al., 2003), the latter mechanism might be occurring in the defined regions of the core variable genome responsible for speciation (Takeuchi et al., 2005; Dordet-Frisoni et al., 2007; Lindsay et al., 2006) such as the coding repeat region of *clfA*, and has been suggested as the basis for *S. aureus* evolution (Kuhn et al., 2006).

A general concordance was obtained by *clfA* and other methods in regional distributions of genotypes and host-specificity. The dominant isolates in each region mostly clustered together and showed similar repeat deletion patterns across the sequences. In agreement with Highlander et al., (2007), the sequence alignment pattern revealed that the USA300 subsets FPR and TCH were closely related to COL and Newman. This is further confirmed by the phylogenetic relationships of isolates in Figure 4.3. A recent study also confirmed the above relationships and further revealed the relative evolutionary distance between the above group and the CA- MSSA strain MSSA476 within the cluster (Baba et al., 2008). Interestingly, this is also consistent with *clfA* typing; all four strains were identical based on copy-number (51 copies) except the MSSA476 (49 copies). Altogether, these findings further confirm the view that CA-MRSA lineages are distinct and may descend from compatible endemic methicillin-susceptible *S. aureus* (MSSA), and not from the hospital-acquired (HA-MRSA) lineages (Groom et al., 2001; Naimi et al., 2001;59 Vandenesch et al., 2003). The similarity of repeat profiles of the above strains might be due to their geographical proximity as follows; COL (Colindate, England), Newman (Hampshire UK), MSSA476 (UK), and the USA300 (first isolated in Denmark in 2000, Larsen et al., 2007) and then in Europe, USA, Canada, and worldwide, although this assumption would

require a comprehensive screening to confirm. In fact, typing by another repeat-based method, the coagulase gene, was useful in revealing the predominant *S. aureus* type in different geographical regions (Saei et al, 2009, Su et al., 1999). Thus, geographic variation in genotypes often correlates with virulence properties. For instance, in dairy cows, herd- and environment-specific subtypes showed significantly differing epidemicity and virulence in different regions (Sommerhauser et al., 2003; Joo et al, 2001; Graber et al., 2009), where the predominant types had higher virulence and fitness (Su et al., 1999; Fitzgerald et al, 1997).

Typing with core variable genes responsible for adaptations such as the coding tandem-repeats (Verstrepen et al., 2005), would potentially help close the gap between clonality and evolution of host-specific hyper-virulent phenotypes. *clfA* repeats are highly degenerate, and showed codon triplet position bias, between human and animal isolates in individual repeat units. The majority of the point mutations detected were transversions (Figure 4.2A). Further, the high sequence-pattern similarity in related groups, prevalence of full repeat unit deletions in the centre of the sequence, and occurrence of a few incomplete repeat-units, would imply a strong positive selection at the "full-repeat-level" and not at the point mutation level to change amino acids individually, supporting the notion that tandem-repeats mediate adaptation by changing the repeat "lengths". This shows the importance of repeat-size over point mutation (Verstrepen et al., 2005). This is further supported by the apparent sequence conservation of nucleotide repeat "units" based on several facts: firstly codon bias is usually significantly higher at conserved rather than at non-conserved positions. In this case, the highly repeated serine codon-triplet is the primordial one (TCN in the first, third, and fifth positions of the 18bp unit). Secondly the serine major codon for conserved residues is TCC (Figure 4.2B), and thirdly serine codon for non-conserved residues is AGC (Diaz-Lazcoz et al., 1995), and thus the latter (5^{th} and 6^{th} triplets) may be the "hot-spot" for potential recombination.

The *spa*-typing revealed seven lineage groups for the 25 PFGE pattern types; four types identified in PFGE lineage group A and three in lineage group D. In group A, each major *clfA* type was mostly associated with one *spa* type. Given the temporal and regional diversity of the isolates, this suggests that in mastitis primary differentiation of *spa* is followed by polymorphism of *clfA*. It follows that the inability of SpA to interact with cow's serum (2 Atkins et al., 2008) might be compensated by the ClfA's antiphagocytic property which parallels that of SpA (Foster, 2005; Higgins et al., 2006); particularly the former was more prevalent in strains tested than the

first. Finally, similar to the *clfA*, certain *spa* types, particularly the predominant type t529/*spa*102, were present in different lineage groups as determined by other methods. In comparison to the Canadian database for MRSA strains, types t267, t359 and t521 were associated with ST97; the last two belonged to D group and the first to A (corresponding RTs, X and C) indicating diversifications within common types (Sung et al., 2008).

The results of the antimicrobial susceptibility testing in the present study as compared with earlier findings of *S. aureus* from dairy herds in Ontario and Quebec provinces (Sabour et al., 2004), reflects a significant change in resistance patterns after a single decade. None of the 87 *S. aureus* isolates used in this study were MRSAs and only three isolates, all from a single farm in Ontario province, were penicillin resistant (CO12, CO14, and CO13) (Table 4.4). Perhaps these could be the same strains; however, the last isolate had a distant PFGE pattern from the other two, suggesting that these isolates may have ancestral differences. In addition, these three resistances appeared to be due to beta-lactamase production rather than development of small colony phenotypes (Vaudaux et al., 2006) as indicated by their growth properties. There was also a single case of tetracycline resistance and many cases of erythromycin intermediate resistance patterns in Eastern provinces (Table 4.4). These coupled with negative result of multiplex PCR amplifications for *mecA* and *pvl* genes, as well as possession of beta type hemolysis are properties resembling that of MSSA phenotypes and mastitis-specialized lineage. This is consistent with the finding of a highly significant regional variation (2% to 65%) in superantigen toxin gene prevalence in mastitis isolates from different countries suggesting their limited role in the pathogenesis of bovine mastitis, in contrast to uniform prevalence of beta hemolysin showing its role (Larsen et al., 2002).

In summary, we found good correlation and comparable discriminatory powers between *clfA* repeat typing and that of PFGE and *spa* typing. Our results were in agreement with others in overall strain clustering and in the clonality of mastitis lineage as shown by *clfA* and other techniques used. In addition, we have found that the distribution and diversity of mastitis-specific *S. aureus* genotypes across Canada was unique and was region specific. Two major PFGE lineage groups A and D mainly represented Eastern and Western Canada strain types, respectively. These groups, 82% of the first and 43% of the latter, were further differentiated by the same four major *clfA* subtypes. Thus, we have further provided a new prospective on the possible recombinogenic nature of the *clfA* locus in a clonal background of the species implying

selective evolution, potentially by recombination, into subclonal populations in different geographical regions across Canada. These would show the usefulness of *clfA* in typing and tracing *S. aureus* strains, and coupling straightforward epidemiological investigations with pathogenesis and colonization potentials. Concordance between copy number typing and whole R domain sequence indicates sequence conservation within repeat units suggesting that the former can be used for screening purposes even without extensive sequencing. Finally, we suggest that it should be considered for strains from hoofed animals, including cows, where SpA interaction is not clear.

ACKNOWLEDGEMENT

We are grateful to Dr. Michel Couillard and the PFGE staff at the LSPQ particularly Joanne Prévost, Guylaine Aubin and Isabelle Robillard, for the excellent collaboration and the opportunity for training on the Canadian Standards of PFGE protocol, to Brigitte Lefebvre and Luc Massicotte and their staff for the antibiotic susceptibility testing of isolates, to Louis Bryden and members of Antimicrobial Resistance and Nosocomial Infections program, at the National Microbiology Laboratory for the training and expert advices on the Bionumerics analysis of the PFGE data and *spa* typing without whom the work would have not seen the light. This study was supported by funding from a discovery grant from Natural Science and Engineering Research Council of Canada and from the Canadian Mastitis Research Network. Additional support was provided by a grant from the National Nature Science Foundation of China (NSFC Grant No. 30828026).

REFERENCES

Ahmed S, Meghji S, Williams R, Henderson B, Brock J, Nair S. 2001. *Staphylococcus aureus* fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria. Infect Immun 69: 2872-2877.

Asao T, Kumeda Y, Kawai T, Shibata T, Oda H, Haruki K, Nakazawa H, Kozaki S. 2003. An extensive outbreak of staphylococcal food poisoning due to low-fat milk in Japan: estimation of enterotoxin A in the incriminated milk and powdered skim milk. Epidemiol Infect 130:33-40.

Atkins K, Burman J, Chamberlain E, Cooper J, Poutrel B, Bagby S, Jenkins A, Feil E, van den Elsen J. 2008. *S. aureus* IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. Mol Immunol 45: 1600-1611.

Baba T, Bae T, Schneewind O, Takeuchi F, Hiramatsu K. 2008. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. J Bacteriol 190: 300-310.

Cheng A, Kim H, Burts M, Krausz T, Schneewind O, Missiakas D. 2009. Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. FASEB J.

Diaz-Lazcoz Y, Hénaut A, Vigier P, Risler J. 1995. Differential codon usage for conserved amino acids: evidence that the serine codons TCN were primordial. J Mol Biol 250: 123-127.

Dordet-Frisoni E, Dorchies G, De Araujo C, Talon R, Leroy S. 2007. Genomic diversity in *Staphylococcus xylosus*. Appl Environ Microbiol 73: 7199-7209.

Enright M, Robinson D, Randle G, Feil E, Grundmann H, Spratt B. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci U S A 99: 7687-7692.

Feil E, Li B, Aanensen D, Hanage W, Spratt B. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186: 1518-1530.

Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt BG, Moore CE, and Day NP. 2003. How clonal is *Staphylococcus aureus*?. J Bacteriol 185: 3307–3316.

Fitzgerald J, Meaney W, Hartigan P, Smyth C, Kapur V. 1997. Fine-structure molecular epidemiological analysis of *Staphylococcus aureus* recovered from cows. Epidemiol Infect 119: 261-269.

Foster T. 2005. Immune evasion by staphylococci. Nat Rev Microbiol 3: 948-958.

Foster T, Höök M. 1998. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 6: 484-488.

Golding G, Campbell J, Spreitzer D, Veyhl J, Surynicz K, Simor A, Mulvey M. 2008. A preliminary guideline for the assignment of methicillin-resistant *Staphylococcus aureus* to a Canadian pulsed-field gel electrophoresis epidemic type using spa typing. Can J Infect Dis Med Microbiol 19: 273-281.

Graber H, Naskova J, Studer E, Kaufmann T, Kirchhofer M, Brechbühl M, Schaeren W, Steiner A, Fournier C. 2009. Mastitis-related subtypes of bovine *Staphylococcus aureus* are characterized by different clinical properties. J Dairy Sci 92: 1442-1451.

Groom A, Wolsey D, Naimi T, Smith K, Johnson S, Boxrud D, Moore K, Cheek J. 2001. Community-acquired methicillin-resistant *Staphylococcus aureus* in a rural American Indian community. JAMA 286: 1201-1205.

Guinane C, Sturdevant D, Herron-Olson L, Otto M, Smyth D, Villaruz A, Kapur V, Hartigan P, Smyth C, Fitzgerald J. 2008. Pathogenomic analysis of the common bovine *Staphylococcus aureus* clone (ET3): emergence of a virulent subtype with potential risk to public health. J Infect Dis 197: 205-213.

Harmsen D, Claus H, Witte W, Rothgänger J, Turnwald D, Vogel U. 2003. Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for spa repeat determination and database management. J Clin Microbiol 41: 5442-5448.

Hartford O, Francois P, Vaudaux P, Foster T. 1997. The dipeptide repeat region of the fibrinogen-binding protein (clumping factor) is required for functional expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. Mol Microbiol 25: 1065-1076.

Higgins J, Loughman A, van Kessel K, van Strijp J, Foster T. 2006. Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. FEMS Microbiol Lett 258: 290-296.

Highlander SK, Hultén KG, Qin X, Jiang H, Yerrapragada S, Mason EO Jr, Shang Y, Williams TM, Fortunov RM, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox GE, Cardenas AM, Muzny DM, Hemphill L, Ding Y, Dugan S, Blyth PR, Buhay CJ, Dinh HH, Hawes AC, Holder M, Kovar CL, Lee SL, Liu W, Nazareth LV, Wang Q, Zhou J, Kaplan SL, Weinstock GM.2007. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. BMC Microbiol 7: 99.

Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabbinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci U S A 101: 9786-9791.

Hunter P. 1990. Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol 28: 1903-1905.

Joo Y, Fox L, Davis W, Bohach G, Park Y. 2001. *Staphylococcus aureus* associated with mammary glands of cows: genotyping to distinguish different strains among herds. Vet Microbiol 80: 131-138.

Josefsson E, Kubica M, Mydel P, Potempa J, Tarkowski A. 2008. In vivo sortase A and clumping factor A mRNA expression during *Staphylococcus aureus* infection. Microb Pathog 44: 103-110.

Juhász-Kaszanyitzky E, Jánosi S, Somogyi P, Dán A, van der Graaf-van Bloois L, van Duijkeren E, Wagenaar J. 2007. MRSA transmission between cows and humans. Emerg Infect Dis 13: 630-632.

Kapur V, Sischo W, Greer R, Whittam T, Musser J. 1995. Molecular population genetic analysis of *Staphylococcus aureus* recovered from cows. J Clin Microbiol 33: 376-380.

Kennedy AD, Otto M, Braughton KR, Whitney A R, Chen L, Mathema B, Mediavilla JR, Byrne KA, Parkins LD, Tenover F C, Kreiswirth BN, Musser JM, and DeLeo F R. 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. Proc Natl Acad Sci U S A 105: 1327-1332.

Klevens R, et al. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. JAMA 298: 1763-1771.

Koreen L, Ramaswamy S, Graviss E, Naidich S, Musser J, Kreiswirth B. 2004. Spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol 42: 792-799.

Koreen L, Ramaswamy S, Naidich S, Koreen I, Graff G, Graviss E, Kreiswirth B. 2005. Comparative sequencing of the serine-aspartate repeat-encoding region of the clumping factor B gene (clfB) for resolution within clonal groups of *Staphylococcus aureus*. J Clin Microbiol 43: 3985-3994.

Kuhn G, Francioli P, Blanc D. 2006. Evidence for clonal evolution among highly polymorphic genes in methicillin-resistant *Staphylococcus aureus*. J Bacteriol 188: 169-178.

Larsen A, Stegger M, Goering R, Sorum M, Skov R. 2007. Emergence and dissemination of the methicillin resistant *Staphylococcus aureus* USA300 clone in Denmark (2000-2005). Euro Surveill 12.

Larsen H, Aarestrup F, Jensen N. 2002. Geographical variation in the presence of genes encoding superantigenic exotoxins and beta-hemolysin among *Staphylococcus aureus* isolated from bovine mastitis in Europe and USA. Vet Microbiol 85: 61-67.

Lindsay J, Moore C, Day N, Peacock S, Witney A, Stabler R, Husain S, Butcher P, Hinds J. 2006. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J Bacteriol 188: 669-676.

McAleese F, Walsh E, Sieprawska M, Potempa J, Foster T. 2001. Loss of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation of transcription, shedding and cleavage by metalloprotease. J Biol Chem 276: 29969-29978.

McDonald R, Antonishyn N, Hansen T, Snook L, Nagle E, Mulvey M, Levett P, Horsman G. 2005. Development of a triplex real-time PCR assay for detection of Panton-Valentine leukocidin toxin genes in clinical isolates of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 43: 6147-6149.

Melles D, et al. 2004. Natural population dynamics and expansion of pathogenic clones of *Staphylococcus aureus*. J Clin Invest 114: 1732-1740.

Mulvey M, Chui L, Ismail J, Louie L, Murphy C, Chang N, Alfa M. 2001. Development of a Canadian standardized protocol for subtyping methicillin-resistant *Staphylococcus aureus* using pulsed-field gel electrophoresis. J Clin Microbiol 39: 3481-3485.

Naimi T, et al. 2001. Epidemiology and clonality of community-acquired methicillin-resistant *Staphylococcus aureus* in Minnesota, 1996-1998. Clin Infect Dis 33: 990-996.

Nanra J, et al. 2009. Heterogeneous in vivo expression of clumping factor A and capsular polysaccharide by *Staphylococcus aureus*: implications for vaccine design. Vaccine 27: 3276-3280.

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

Palmqvist N, Josefsson E, Tarkowski A. 2004. Clumping factor A-mediated virulence during *Staphylococcus aureus* infection is retained despite fibrinogen depletion. Microbes Infect 6: 196-201.

Projan SJ, and RP Novick. 1997. The molecular basis of pathogenicity, p. 55-82. *In* K. B. Crossley and G. L. Archer (ed.), The staphylococci in human disease. Churchill Livingstone, New York, N.Y.

Risley A, Loughman A, Cywes-Bentley C, Foster T, Lee J. 2007. Capsular polysaccharide masks clumping factor A-mediated adherence of *Staphylococcus aureus* to fibrinogen and platelets. J Infect Dis 196: 919-927.

Robinson D, Enright M. 2004. Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. Clin Microbiol Infect 10: 92-97.

Sabour P, Gill J, Lepp D, Pacan J, Ahmed R, Dingwell R, Leslie K. 2004. Molecular typing and distribution of *Staphylococcus aureus* isolates in Eastern Canadian dairy herds. J Clin Microbiol 42: 3449-3455.

Saei H, Ahmadi M, Mardani K, Batavani R. 2009. Molecular typing of *Staphylococcus aureus* isolated from bovine mastitis based on polymorphism of the coagulase gene in the north west of Iran. Vet Microbiol 137: 202-206.

Said K, Ramotar K, Zhu G, Zhao X. 2009. Repeat-based subtyping and grouping of *Staphylococcus aureus* from human infections and bovine mastitis using the R-domain of the clumping factor A gene. Diagn Microbiol Infect Dis 63: 24-37.

Seegers H, Fourichon C, Beaudeau F. 203. Production effects related to mastitis and mastitis economics in dairy cattle herds. Vet Res 34: 475-491.

Sommerhäuser J, Kloppert B, Wolter W, Zschöck M, Sobiraj A, Failing K. 2003. The epidemiology of *Staphylococcus aureus* infections from subclinical mastitis in dairy cows during a control programme. Vet Microbiol 96: 91-102.

Stranger-Jones Y, Bae T, Schneewind O. 2006. Vaccine assembly from surface proteins of *Staphylococcus aureus*. Proc Natl Acad Sci U S A 103: 16942-16947.

Su C, Herbelin C, Frieze N, Skardova O, Sordillo L. 1999. Coagulase gene polymorphism of *Staphylococcus aureus* isolates from dairy cattle in different geographical areas. Epidemiol Infect 122: 329-336.

Sung J, Lindsay J. 2007. *Staphylococcus aureus* strains that are hypersusceptible to resistance gene transfer from enterococci. Antimicrob Agents Chemother 51: 2189-2191.

Sung J, Lloyd D, Lindsay J. 2008. *Staphylococcus aureus* host specificity: comparative genomics of human versus animal isolates by multi-strain microarray. Microbiology 154: 1949-1959.

Takeuchi F, et al. 2005. Whole-genome sequencing of staphylococcus haemolyticus uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. J Bacteriol 187: 7292-7308.

Tenover F, Arbeit R, Goering R, Mickelsen P, Murray B, Persing D, Swaminathan B. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33: 2233-2239.

Tuchscherr L, Buzzola F, Alvarez L, Lee J, Sordelli D. 2008. Antibodies to capsular polysaccharide and clumping factor A prevent mastitis and the emergence of unencapsulated and small-colony variants of *Staphylococcus aureus* in mice. Infect Immun 76: 5738-5744.

van Belkum A. 2007. Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). FEMS Immunol Med Microbiol 49: 22-27.

van Leeuwen W, et al. 2005. Host- and tissue-specific pathogenic traits of *Staphylococcus aureus*. J Bacteriol 187: 4584-4591.

van Loo I, Huijsdens X, Tiemersma E, de Neeling A, van de Sande-Bruinsma N, Beaujean D, Voss A, Kluytmans J. 2007. Emergence of methicillin-resistant *Staphylococcus aureus* of animal origin in humans. Emerg Infect Dis 13: 1834-1839.

Vandenesch F, et al. 2003. Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. Emerg Infect Dis 9: 978-984.

Vaudaux P, Kelley W, Lew D. 2006. *Staphylococcus aureus* small colony variants: difficult to diagnose and difficult to treat. Clin Infect Dis 43: 968-970.

Verstrepen K, Jansen A, Lewitter F, Fink G. 2005. Intragenic tandem repeats generate functional variability. Nat Genet 37: 986-990.

Villard L, Lamprell H, Borges E, Maurin F, Noël Y, Beuvier E, Chambaand JF, Kodjo A. 2005. Enterotoxin D producing strains of *Staphylococcus aureus* are typeable by pulsed-field gel electrophoresis (PFGE). Food Microbiol 22:261-265.

Wolz C, McDevitt D, Foster T, Cheung A. 1996. Influence of agr on fibrinogen binding in *Staphylococcus aureus* Newman. Infect Immun 64: 3142-3147.

CONNECTING STATEMENT 3

In previous chapters we have developed and evaluated the repeat-based sub-typing marker of *clfA*, investigated its capability to couple epidemiological typing with host- and organ-specific colonization potentials. These chapters have lead to precise identification of major clones of *S. aureus* bovine-mastitis specialized lineages with common properties. In Chapter V, we used a representative of the major clone (strain764) for invasion into bovine mammary epithelial cells. A high-throughput whole-genome transcriptomic analysis of the internalized bacteria would lead to identification of subsets of genes and regulatory pathways that were specifically expressed intracellular in the bacteria within the mammary epithelial cells.

CHAPTER V: GENOME-WIDE COMPARATIVE EXPRESSION PROFILES OF HUMAN- AND MASTITIS-SPECIFIC STAPHYLOCOCCUS AUREUS LINEAGES AFTER THEIR INTERNALIZATION INTO MAMMARY CELLS

Kamaleldin B Said[1], Marcus B Jones[2], Alexander I Saeed[2], Tatiana Dracheva[2], Scott Peterson[2] and Xin Zhao[1,*].

[1]McGill University, Department of Animal Science**,** Room MS1-084, Macdonald Stewart Building, 21111 Lakeshore Road, Ste. Anne de Bellevue, Quebec H9X 3V9.

[2]J. Craig Venter Institute (JCVI), Pathogen Functional Genomics Resource Center, 9704 Medical Center Drive, Rockville, MD 20850, USA

[1]*Xin Zhao, corresponding author. Mailing address: Department of Animal Science**,** Room MS1-084, Macdonald Stewart Building, 21111 Lakeshore Road, Ste. Anne de Bellevue, Quebec H9X 3V9. Telephone:514-398-7975, Fax: 514-398-7964

Ready for submission

ABSTRACT

Staphylococcus aureus is an important re-emerging pathogen in humans and major food animals. Although the genome is conserved, how separate lineages adapt to and cause humans infection and bovine mastitis has been quite elusive. We have used mammary epithelial cell internalized human and bovine isolates for genome-wide expression to study the impact of oxygen, which is low in mammary gland, and to identify subsets of genes expressed in the mammary, using high-throughput genome qRT-PCR. In the mastitis isolate under normal oxygen, expression of transcriptional regulators *MerR, sigB*, VraS, YycG/YycF, araC, and *tetR* indicated adaptation, environmental sensing, binding, and protection. Coupling of fermentative metabolism to virulence was evident by upregulation of the catabolite control protein A (*ccpA)*, pentose monophosphate pathway, and down regulation of TCA cycle intermediates. Invasive phenotypes, potentially through *sarU* activation of *agr* and expression of toxins, adaptation, detoxification, and *in vivo* viability factors as staphopains and GntR operon, also occurred. However, under reduced oxygen fibrinogen binding, *isd* operon, (*isdA, isdC, isdH)*, and *sdrH* were experessed, indicating aggressive binding phenotype. Irrespective of oxygen condition, strain Wright was less aggressive showing intense intracellular signalling potentially through *MerR* and *sigB*, limited toxins, and many hypothetical genes. Interestingly, under reduced oxygen, universal down regulation in metabolic and regulatory pathways occurred, showing persistence of the Wright strain. Taken together, the quick response of the mastitis isolate strongly suggested adapted phenotype to mammary cells. In addition, reduced oxygen significantly altered expresson patterns, with the mastitis isolate inducing more persistence properties. Future annotation the large number of hypothetical proteins would potentially reveal a more precise network of regulatory subsets that are specifically expressed in the mammary gland.

INTRODUCTION

The increasing availability of whole genome sequences has improved our understanding on the basic mechanisms underlying virulence, pathogenicity, and specialization of major human and animal pathogens, such as *S. aureus*. Due to its importance in both medical and veterinary medicine, *S. aureus* has been one of the focuses in present day research. Over 14 genome sequences have been completed for the species to date. Despite enormous efforts, however, how is *S. aureus* emerged as a successful pathogen has been quite elusive.

Differential expression of chromosome encoded genes has increasingly been found to enhance virulence. Much less has been documented on this about mastitis strains in comparison with human strains. Recently, it was reported that differential expression of core genome encoded key factors such as the phenol soluble modulins and alpha toxins was responsible for evolution of the currently emerging strains (Li et al., 2009). Expression patterns of these factors were same in MSSA and MRSA isolates of the same type consistent with the finding that methicillin-resistence, although it may aid in transmission, does not have an impact on virulence per se nor does it affect strain fitness (Bubeck Wardenburg et al., 2008; Li et al., 2009). Furthermore, pheno- and genotypically identical strains may differ in their acquisition of the panton-valentine leukocidin gene (Zhang et al., 2008). Thus, the notion that the transmissible genetic elements are always responsible for *S. aureus* virulence has been changing in recent years (Rossney et al., 2007). The above findings indicate that differential expression is a key for virulence pathways, even at the level of single loci. For instance, expression patterns of surface clumping factor A (*clfA*) and *clfB* make them ideal tools for adaptation. These two proteins have similar structural and functional properties; however, they operate under different conditions. The ClfB is expressed at the log phase and increased by oxygenation but it is limited at the stationary phase (McAleese *et al*., 2001; Ni Eidhin et al., 1998). On the other hand, significant ClfA activity has been demonstrated in the stationary phase by intra-mammary gland (IMI) challenge of mice immunized with anti-ClfA antibody (Tuchscherr et al., 2005). Similarly, enhanced *clfA* and *spa* (another tandem repeat containing gene codes for protein A) expression was detected under bovine-mammary

mimicked low oxygen condition, irrespective of strains used (Ster et al., 2005). Using *clfA* repeats, we have recently provided evidence for existance of subclonal populations of *S. aureus* within major PFGE and *spa* groups, in different geographic regions in Canada (Said et al., 2010).

Invasion of *S. aureus* into bovine mammary epithelial cell line (Bayles et al., 1998) has been shown to be controlled by the Agr and Sar systems, and that persistence of *S. aureus* in the epithelial cells required metabolically inactive bacteria (Wesson et al., 1998). This was supported by the finding of Senn et al., (2005) that the small colony variants were successful persistents due to their metabolic dormancy that is partially regulated by *sigB* and its dependent transcripts such as *sarC* and *clfA*. However, attempts to study single regulator mutants could result in several alterations in expression pattern of global additional regulators, such as *arl*, *rot*, *sae*, *sarR*, *sarS*, *srr*, *svrA*, making it difficult to determine the effect of a single gene product on each pathway (Senn et al., 2005). Thus, it is not clear which pathways are induced in the mammary gland. Furthermore, use of laboratory passed strains may not be suitable due to mutations (Somerville et al., 2002; Wertheim et al., 2008; Grundmeier et al., 2004).). Another factor is the mutation in the regulatory gene itself. One example is an *agr*-mediated bacterial interference in which, due to sequence variation in *agrB*, a*grD*, and *agrC*, AIP produced by one strain inhibits the function of other *agr* types (Novick, 2003). In addition, the genetic background of the host also plays important roles in the determination of resident strains (Bubeck Wardenburg et al., 2008). The environment of the host also affected gene expression of resident strains. A recent study used low oxygen levels in a fermentor condition to mimic the bovine mammary gland for gene expression by RealTime PCR. Certain genes were not affected by oxygen level or strain, such as the regulatory factors RNAIII, *rot*, and *sarR* genes, and the virulence factors *asp23* and *clfA* genes, and were strongly expressed. However, under low oxygenation, *srr* and *spa* genes were expressed. Expression of some regulators such as *sae, sarA* and *sigB* was affected by the strain and the oxygenation condition (Ster et al., 2005). For these inconsistencies, existence of unidentified alternative pathways was suggested by the authors. Thus, the present study investigates the whole-genome expression profiles of mastitis specific *S. aureus* isolates,

that were directly isolated from clinical conditions, during invasion into mammary epithelial cells. As a comparison, a human strain, Wright, was also included.

MATERIALS AND METHODS

Bacterial strains

A *S. aureus* isolate (mastitis-associated isolate 746) which represented a dominant clone that we have recently identified during a regional profiling (PFGE A) and included a subtype (PFGE A7) corresponded to *clfA* subtype Q identical to the sequenced RF122 strain (Said et al., 2010) was used for invasion of mammary epithelial cells (MAC-T) grown under the normal (with 5% CO_2) and reduced oxygen conditions. A human associated *S. aureus* ATCC49525 strain (dominant isolate from human infections) (designated hereafter as strain Wright) was also used under the same conditions.

S. aureus invasion into mammary epithelial cell line (MAC-T cells)

Preparation of bacterial cultures for invasion experiment was carried out by combination of methods from Bayles et al. (1998) and Wesson et al. (1998) with some modifications. The overnight cultures (grown in the invasion medium, see below) were centrifuged and the pellet was washed only once with sterile phosphate-buffered saline (pH 7.2) and resuspended in 10 ml of the invasion medium (see below) to give a cell density of 10^{10} CFU ml^{-1}. Serial dilutions of this were prepared and 1 ml/well of the 10^2 dilution was used to inoculate MAC-T cell monolayers at the multiplicity of infections of (MOI 100). *S. aureus* mastitis isolate 764 and the human associated strain Wright were used for invasion of MAC-T mammary epithealial cells grown at normal and reduced oxygen conditions. Cell culture incubators with 5%CO were used for normal oxygen conditions. For reduced oxygen, ready gas mixtures containing 5% oxygen and balanced by nitrogen were used. The Modular Incubator Chamber (Billups-Rothenberg, Inc, California, USA) was specifically used for reduced oxygen condition. It was first flushed for 15-20 min with the gas mixture after loading with co-cultures, closed gas tight, and connecting to the gas cylinders though specialized tubings and valves.

Cell culture

An established bovine mammary epithelial cell line, designated MAC-T (Huynh et al., 1991) was used for invasion experiments. Composition of the MAC-T cell growth medium was 44.5% Dulbecco's Modified Eagle Medium (Gibco BRL), 44.5%RPMI Medium 1640 (1X) with L-glutamine (Sigma-Aldrich Ltd, Oakville, Ontario, Canada), 10% Fetal Bovine Serum (FBS) (Invitrogen Canada Inc., Burlington, Ontario, Canada), and 1ml of 100x concentrated antibiotic/antimycotic solution (Invitrogen Inc.), sterilized by filtration using 0.22 μm pore size. Cells were seeded at 1 x 10^6 cells/well and grown for approximately 2-3 days at 37°C with 5% CO_2. For reduced oxygen environment, Modular Incubator Chamber model with oxygen level adjusted to 5% was used to incubate cell culture plates (Invasion co-cultures) as well as separate bacterial cultures (non-reacted controls). Cells were grown in 6-well tissue culture plates for invasion.

Invasion Assay

The invasion was carried out as described by Shompole et al. (2003). Approximately 16 h prior to invasion experiment the MAC-T cell growth medium was replaced with 1 ml of invasion medium (growth medium without antibiotics and FBS). The morning of the experiment, the medium was removed and MAC-T cells were washed once with the invasion medium. Appropriate wells of MAC-T cells were then inoculated with 1ml of the 10^2 diluted *S. aureus* and incubated at 37°C with 5% CO_2.To synchronize invasion process plates were immediately shaken in swirling motion manually for a few min. After 1 h, supernatants of the cocultures were removed and monolayers were then washed three times with lysostaphin (10 μg ml^{-1}; Sigma) in invasion medium at 37°C to kill extracellular bacteria (Qazi et al., 2004; Garnozi et al., 2007) before incubating in fresh invasion medium containing 100ug/ml gentamicin (Invitrogen Inc.) for 6-8 h. Non-reacted bacterial cultures were grown under the same condition as negative controls. Supernatants were then removed and discarded. MAC-T cell monolayers were carefully and quickly washed with sterile distilled water, scraped from plates with the help of disposable sterile cell scrapers (Fisher Scientific, Ottawa, Ontario, Canada), and quickly stored at -80 °C in 5-10 volumes of RNAlater solution as described by the manufacturer (Applied Biosystems, Ambion Inc., Canada, Streetsville, Ontario, Canada), before RNA

extractions after homogenization by bead beating systems for 45 sec at max speed according to the Ambion protocol for RiboPure Bacteria kit (Ambion Inc.) at the PFGRC, J. Craig Venter Institute (JCVI), Maryland, USA.

Whole-genome expression profiling of intracellular S. aureus

cDNA synthesis and purification

cDNA was synthesize by taking 2 µg of total RNA, mixing with a final concentration of 0.5 µM of dNTP mix and random hexamers. cDNA was purified using Qiagen minElute column (Qiagen), according to the manufacturer protocol.

High-throughput qRT-PCR

Purified cDNA was diluted 1:10 in DEPC water. 9ml of diluted cDNA and 15ml of Roche 2x SYBR Green master were mixed. 8 µl of cDNA/mastermix was aliquoted into forty 384-well plates. Forward and reverse primers (whole-genome primers were made on the complete sequences of *S. aureus* strains by the JCVI, Maryland) were added to cDNA master mix to a final volume of 10µl. Plates were sealed with qRT-PCR tape and stored in the -80 until use. Plates were run in a Roche LifghtCycler480 (LC480) RealTime PCR system 384-plate format for gene-expression, interfaced with robotics to create an automated high-throughput. qRT-PCR cycles were as follow: Step 1: 5 minutes at 95C. Step 2 (65 cycles): 95 degree for 10 sec, 60 degree for 10 sec, 72 degree to 10 secs. Step 3: Melt curve analysis on all wells.

Normalization and data analysis

The data from the qRT-PCR was normalized (shifted) on the average Cp value by the plate. Each sample was done in duplicates on plates to find the average cp values for measurements, and the duplicates (on the same plate) were averaged before normalization. Experimental data were then were then normalized based on relative expression levels between each experiment and its control under the given experimental condition. If only one value from the duplicate was available, it was excluded from the analysis. The difference between the Cp values between the experiment and control

(CpExperiment – CpControl) were considered for measurements of expression profiles across the genome. The difference was considered significant only if it was higher than the difference between duplicates in both samples and more than 2 (4 fold difference). In other words, normalizations were carried out by first getting the plate mean which equals sum of all cp (crossing points)/384 (number of well in each plate). Then by substracting that mean from cp1 to cp384 (treatment samples) and (control samples). This difference in expression was considered if it was high (at least 4 fold).

Biological role query tool (BioQT v1.1.0)

The expression data generated by the high-throughput LC480 system was run in the Biological Role Query Tool (BioQT) linked to the server at the JCVI. BioQT retrieves and presents annotation information from several databases (CMR Cellular Roles, Gene Ontology (GO) Terms, Protein Families (PFam), TIGRFams, Kyoto Encyclopedia for Genes and Genomes (KEGG) Pathways, Enzyme Commission information (EC) numbers, and KEGG Orthologs for lists of protein accessions or gene locus IDs for all organisms currently entered into JCVI's Comprehensive Microbial Resource (CMR) or into a JCVI small genome database. Results are presented graphically in the result viewer and can be output to tab delimited files. The distribution of cellular roles found within lists of genes or proteins in the searched databases can provide high level insight into biological systems that are influenced by the conditions under study. BioQT reports summaries of cellular roles and also includes a pie chart that represents the role distribution. The chart is customizable, uses the standard CMR role color scheme, and provides options for direct data output and chart image output. There are five types of file output generated. The database search is given as a file extension name. 1) ROLE file (CMR Cellular Roles), 2) GO file (GO ID and Term file), 3) PFAM file (HMM Protein Family File), 4) TIGRFAM file [(HMM **(**hidden Markov model) Protein Family File], and 5) ROLE_SUM file (CMR Cellular Role Summary File). The column heading in all five files are almost identical except for database name columns.

RESULTS

Genome-wide expression profiles of mastitis *S. aureus* isolate 764 internalized into mammary cells at normal oxygen condition

The whole genome expression profiles of mammary epithelial cell-internalized *S. aureus* were arranged in terms of biological role categories. The data sets obtained from all expression experiments are shown in the corresponding tables and figures. Full data for all sets of experiemtns is available from the supplementary material. As shown in Figure 5.1 and Table 5.1, expressions of specific gene classes were up regulated at normal oxygen condition. Among the most important were those encoding cellular processes, (pathogenesis, adaptation to atypical conditions, and toxin production and resistance), transport and binding proteins, regulatory functions, energy metabolism and a large number of hypothetical proteins. In the pathogenesis category, genes expressed were *agr* signaling molecules, *sarU*, *vraSR*, as well as toxin production including alpha and delta hemolysins, aerolysins/leukocidin family, superantigen-like proteins, and cysteine proteases (Table 5.1). Another major change in gene expression profiles was seen in the regulatory and environmental sensing. For instance, expression of the MerR, which activates sigma factor, and responds to environmental cues, was up regulated. One of the most common DNA-binding transcriptional regulators involved in pathogenesis and environmental stresses response called *araC,* was significantly increased. Furthermore, up regulation of chromosomally encoded transcriptional regulators of drug resistance and fitness, such as GntR, TetR, and MmpL efflux pumps has also occurred (Table 5.1). Similarly, expression of the genome encoded transcriptional regulators of the genes encoding multidrug efflux pumps NorA, NorB, NorC, and AbcA, such as the NorG of the GntR family, that regulates fitness and resistance, were observed. In addition, genes required for adaptation and detoxification were expressed such as TetR, *pnbA* (para-nitrobenzyl esterase) for hydrolysis of beta lactam esters PNB esters, *betA* and *betB* for dehydrogenation of choline and betaine, respectively.

Interestingly, energy metabolism was favorably induced through glycolysis, fermentation, and anaerobic pathways were expressed, as the genes encoding enzymes of these pathways were expressed. In the Glycolysis sub role, genes for fructose-

bisphosphatase, L-lactate dehydrogenase(*ldh1*), and fructose-1,6-bisphosphate aldolase (*fdaB*) were up regulated. Other genes included aldehyde dehydrogenase (*aldA2*), bifunctional acetaldehyde-CoA/alcohol dehydrogenase, and alpha-acetolactate decarboxylase (*budA2*) were expressed, in addition to anaerobic enzymes pyruvate-formate-lyase-activating enzyme (*pflA*) and formate acetyltransferase genes (*pflB*) for fermentation. Finally, increased differential expression of core chromosomal genes of the phenol soluble modulin has occurred during the invasion process. However, the large majority of genes with unknown functions were expressed as shown in the class of hypothetic proteins.

The down regulated genes in isolate 764 under normal oxygen condition are shown in Figure 5.2 and details of main role and sub role are shown in Table 5.2. Metabolism, biosynthesis, cell envelope, and transport and binding protein classes were largely down regulated confirming their limited expression shown in Figure 5.1. Hypothetical proteins that were found mostly in bovine strain RF122 were also down regulated. Similarly, in the "other" section of the regulatory function category, truncated cell surface fibronectin-binding protein was found (locus SAB1289c). Furthermore, PTS system component IIABC (locus SACOL2552) was down regulated. Unique gene categories in biosynthetic and central metabolic pathway were affected such as L-lactate dehydrogenase (*ldh2*) locus (SACOL261). Nevertheless, as shown above, most the virulence associated genes including signal transduction and cellular process categories showed no significant down regulation.

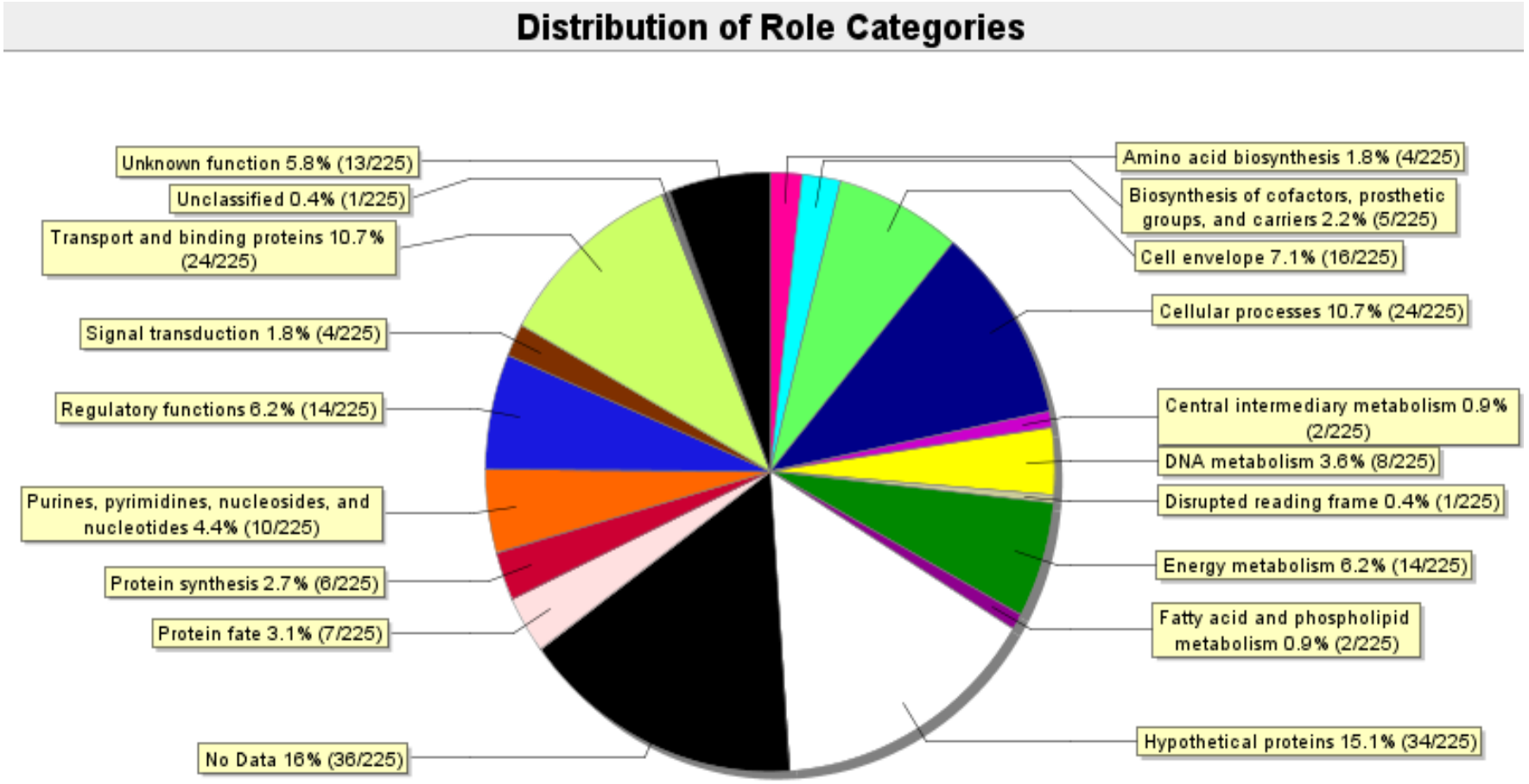


Figure 5.1. Subsets of up regulated gene classes in mastitis *S. aureus* 764 internalized into mammary cell at normal oxygen condition.

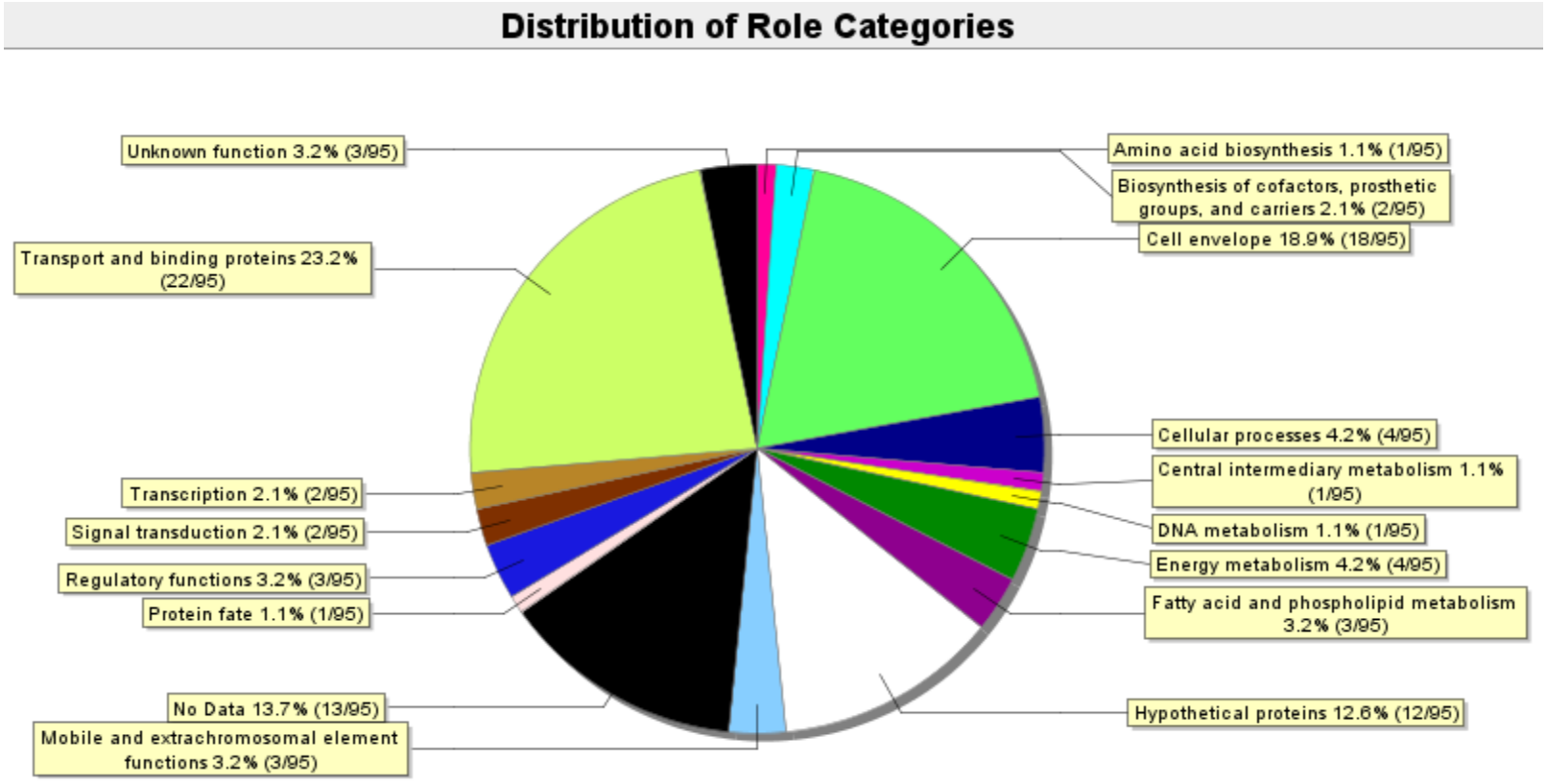


Figure 5.2. Subsets of down regulated gene classes in mastitis *S. aureus* 764 internalized into mammary cells at normal oxygen condition.

Table 5.1. Upregulated genes is mastitis *S. aureus* 764 internalized into mammary cells at normal O_2 condition

Main Role	MC	Sub-role	Product	Locus	Org. Descr.
Cellular processes	24	Detoxification	pnbA, hydrolyse b-lactam PNB	SACOL2459	S.aureus . COL
Cellular processes	24	Adaptations to atypical conditions	betA, choline deHse	SACOL2627	S.aureus . COL
Cellular processes	24	Adaptations to atypical conditions	betB, betaine aldehyde deHase	SACOL2628	S.aureus . COL
Cellular processes	24	Cell division		SACOL1193	S.aureus . COL
Cellular processes	24	Pathogenesis	agrA, response regulator AgrA	SACOL2026	S.aureus . COL
Cellular processes	24	Pathogenesis	agrD	SACOL2024	S.aureus . COL
Cellular processes	24	Pathogenesis	putative	SACOL1184	S.aureus . COL
Cellular processes	24	Pathogenesis	putative	SACOL1184	S.aureus . COL
Cellular processes	24	Pathogenesis	sarU	SACOL2507	S.aureus . COL
Cellular processes	24	Pathogenesis	sarU	SACOL2507	S.aureus . COL
Cellular processes	24	Pathogenesis	sspB2	SACOL1970	S.aureus . COL
Cellular processes	24	Pathogenesis	hld	SACOL2022	S.aureus . COL
Cellular processes	24	Pathogenesis	hld	SACOL2022	S.aureus . COL
Cellular processes	24	Pathogenesis	putative	SACOL2004	S.aureus . COL
Cellular processes	24	Pathogenesis	putative	SACOL2004	S.aureus . COL
Cellular processes	24	Pathogenesis	hlY	SACOL1173	S.aureus . COL
Cellular processes	24	Pathogenesis	hlY	SACOL1173	S.aureus . COL
Cellular processes	24	Pathogenesis	Aerolysis/leukocidin family	SACOL2006	S.aureus . COL
Cellular processes	24	Pathogenesis	Aerolysis/leukocidin family	SACOL2006	S.aureus . COL
Cellular processes	24	Cell adhesion	camS, hypothetical protein	SACOL1964	S.aureus . COL
Cellular processes	24	Cell adhesion	camS, hypothetical protein	SACOL1964	S.aureus . COL
Cellular processes	24	Other	camS, hypothetical protein	SACOL1964	S.aureus . COL
Cellular processes	24	Other	camS, hypothetical protein	SACOL1964	S.aureus . COL
Cellular processes	24	Toxin production and resistance	vraS, sensor histidine kinase VraS	SACOL1943	S.aureus . COL
Cellular processes	24	Toxin production and resistance	superantigen-like protein	SACOL0470	S.aureus . COL
Cellular processes	24	Toxin production and resistance	set11, superantigen-like protein	SAB0376	S.aureus . COL
Cellular processes	24	Toxin production and resistance	exfoliative toxin, putative	SACOL1184	S.aureus . COL
Cellular processes	24	Toxin production and resistance	exfoliative toxin, putative	SACOL1184	S.aureus . COL
Cellular processes	24	Toxin production and resistance	sarU	SACOL2507	S.aureus . COL
Cellular processes	24	Toxin production and resistance	sarU	SACOL2507	S.aureus . COL
Cellular processes	24	Toxin production and resistance	hld	SACOL2022	S.aureus . COL
Cellular processes	24	Toxin production and resistance	hld	SACOL2022	S.aureus . COL
Cellular processes	24	Toxin production and resistance	Leukocidin, putative	SACOL2004	S.aureus . COL
Cellular processes	24	Toxin production and resistance	hlY, a-hemolysin precursor	SACOL1173	S.aureus . COL
Cellular processes	24	Toxin production and resistance	hlY, a-hemolysin precursor	SACOL1173	S.aureus . COL
Cellular processes	24	Toxin production and resistance	Aerolysin/leukocidi family protein	SACOL2006	S.aureus . COL
Cellular processes	24	Toxin production and resistance	Aerolysin/leukocidi family protein	SACOL2006	S.aureus . COL
Transport and binding prot	24	Amino acids, peptides and amines	aa ABC transporter, ATP-binding prot	SACOL2453	S.aureus . COL

Table 5.1 Continued

Transport and bindg prot	24	Amino acids, peptides and amines	Oligo pep, ATP-bindg prot	SAB0861	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	aa ABC transporter, permease	SACOL2452	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	Oligo pep transport system permease	SAB0857	S.aureus .RF122
Transport and bindg prot	24	Amino acids, peptides and amines	Oligo pep, ATP-bindg prot	SAB0855	S.aureus .RF122
Transport and bindg prot	24	Amino acids, peptides and amines	Oligo pep, ATP-bindg prot	SAB0860	S.aureus .RF122
Transport and bindg prot	24	Amino acids, peptides and amines	ABC transporter permease, Putative,	SACOL2475	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	ABC transporter permease, Putative,	SACOL2476	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	aa ABC transporter, ATP-bindg prot	SACOL2473	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	ABC transporter permease,	SACOL2474	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines	aa ABC transporter, ATP-bindg prot	SACOL2472	S.aureus . COL
Transport and bindg prot	24	Amino acids, peptides and amines		SAB0862	S.aureus RF122
Transport and bindg prot	24	Carbohydrates, organic alcohols, and acids	PTS system, IIBC components	SACOL0178	S.aureus . COL
Transport and bindg prot	24	Carbohydrates, organic alcohols, and acids	citrate transporter, permease protein	SACOL2636	S.aureus . COL
Transport and bindg prot	24	Other	multidrug resistance protein	SAB1892c	S.aureus .RF122
Transport and bindg prot	24	Other	MmpL efflux pump, putative	SACOL2566	S.aureus . COL
Transport and bindg prot	24	Nucleosides, purines and pyrimidines	pbuX, xanthine permease	SACOL0459	S.aureus . COL
Transport and bindg prot	24	Nucleosides, purines and pyrimidines	uraA, uracil permease	SACOL1211	S.aureus . COL
Transport and bindg prot	24	Unknown substrate	aa ABC transporter, ATP-bindg prot	SACOL2525	S.aureus . COL
Transport and bindg prot	24	Unknown substrate	Hypothetical protein	SAB0137	S.aureus .RF122
Transport and bindg prot	24	Unknown substrate	transporter, putative	SACOL2471	S.aureus . COL
Transport and bindg prot	24	Cations and iron carrying compounds	copper-translocating P-type ATPase	SACOL2572	S.aureus . COL
Transport and bindg prot	24	Cations and iron carrying compounds	sodium transport family protein	SACOL2011	S.aureus . COL
Transport and bindg prot	24	Cations and iron carrying compounds	iron-binding protein, putative	SACOL0203	S.aureus . COL

Table 5.1 Continued					
Cell envelope	16	Biosyn&degradn of murein sacculus & peptidog	mraY, P-N-acetylmuramoyl-pentapept-transferase	SACOL1195	S.aureus . COL
Cell envelope	16	Biosyn&degradn of murein sacculus & peptidog	murD, UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	SACOL1196	S.aureus . COL
Energy metabolism	14	Glycolysis/gluconeogenesis	fructose-bisphosphatase	SAB2390	S.aureus .RF122
Energy metabolism	14	Glycolysis/gluconeogenesis	ldh1, L-lactate dehydrogenase	SACOL0222	S.aureus . COL
Energy metabolism	14	Glycolysis/gluconeogenesis	ldh1, L-lactate dehydrogenase	SACOL0222	S.aureus . COL
Energy metabolism	14	Glycolysis/gluconeogenesis	fdaB, fructose-1,6-bisp aldolase	SACOL2622	S.aureus . COL
Energy metabolism	14	Fermentation	aldA2, aldehyde dehydrogenase	SACOL1984	S.aureus . COL
Energy metabolism	14	Fermentation	bifunc acetaldehyde-CoA/alco deHase	SACOL0135	S.aureus . COL
Energy metabolism	14	Fermentation	budA2, alpha-acetolactate decooHase	SACOL2617	S.aureus . COL
Energy metabolism	14	Anaerobic	pflA,pyruva formate-lyase-activtig enz	SACOL0205	S.aureus . COL
Energy metabolism	14	Anaerobic	pflB, formate acetyltransferase	SACOL0204	S.aureus . COL
Energy metabolism	14	Anaerobic	ldh1, L-lactate dehydrogenase	SACOL0222	S.aureus . COL
Energy metabolism	14	Anaerobic	L-lactate dehydrogenase	SACOL0222	S.aureus . COL
Energy metabolism	14	Electron transport	acpD, azoreductase	SACOL0190	S.aureus . COL
Energy metabolism	14	Amino acids and amines	sdaAB, L-serine dehydratase	SACOL2545	S.aureus . COL
Energy metabolism	14	Amino acids and amines	1-pyrroline-5-carboxylate dehydrogenase	SACOL2569	S.aureus . COL
Energy metabolism	14	Amino acids and amines	arcC1, carbamate kinase	SACOL1182	S.aureus . COL
Energy metabolism	14	Amino acids and amines	arcB1, ornithine carbamoyltransfase	SACOL1181	S.aureus . COL
Regulatory functions	14	DNA interactions	MerR family transcriptional regulor	SACOL2517	S.aureus . COL
Regulatory functions	14	DNA interactions	AraC family DNA-bindg resp regulator	SACOL0201	S.aureus . COL
Regulatory functions	14	DNA interactions	AraC family DNA-bindg respo regulato	SACOL0201	S.aureus . COL
Regulatory functions	14	DNA interactions	RpiR p-sugar-bindg transcrip regul	SACOL0179	S.aureus . COL
Regulatory functions	14	DNA interactions	sarU, staphyloco accessory regulator U	SACOL2507	S.aureus . COL
Regulatory functions	14	DNA interactions	sarU, staphylococ accessory regulato U	SACOL2507	S.aureus . COL
Regulatory functions	14	DNA interactions	GntR family transcriptional regulor	SACOL0518	S.aureus . COL
Regulatory functions	14	DNA interactions	TetR family transcriptional regulor	SACOL2610	S.aureus . COL
Regulatory functions	14	RNA interactions	pyrR, pyrimidine regulatory prot PyrR	SACOL1210	S.aureus . COL
Regulatory functions	14	Other	agrA,accessory gene regulator prot A	SACOL2026	S.aureus . COL
Regulatory functions	14	Other	agrD, accessory gene regulator prot D	SACOL2024	S.aureus . COL
Regulatory functions	14	Other	sarU, staphylococ accessory regulat U	SACOL2507	S.aureus . COL

Table 5.1 Continued….

Regulatory functions	14	Other	sarU, staphylococ accessory regulat U	SACOL2507	S.aureus . COL
Regulatory functions	14	Protein interactions	vraS, sensor histidine kinase VraS	SACOL1943	S.aureus . COL
Regulatory functions	14	Protein interactions	yycG, sensory box histidine kinase	SACOL0020	S.aureus . COL
Regulatory functions	14	Protein interactions	AraC family DNA-bindg resp reglor	SACOL0201	S.aureus . COL
Regulatory functions	14	Protein interactions	AraC family DNA-bindg resp regulor	SACOL0201	S.aureus . COL
Regulatory functions	14	Protein interactions		SACOL1232	S.aureus . COL
Unknown function	13	Enzymes of unknown specificity	deHdase/reductase family	SACOL2488	S.aureus . COL
Unknown function	13	Enzymes of unknown specificity	exonuclease	SACOL1954	S.aureus . COL
Unknown function	13	Enzymes of unknown specificity	alpha/beta fold family hydrolase	SACOL2597	S.aureus . COL
Unknown function	13	General	hypothetical protein SACOL0014	SACOL0014	S.aureus . COL
Unknown function	13	General	murQ, N-acetylmuramic acid-6-petherase	SACOL0177	S.aureus . COL
Unknown function	13	General	glyoxalase family (detoxification)	SACOL1207	S.aureus . COL
Unknown function	13	General	yycI, hypothetical protein SACOL0022	SACOL0022	S.aureus . COL
Unknown function	13	General	geranylgeranylglyceryl phosphate synthase-like protein	SACOL1967	S.aureus . COL
Unknown function	13	General	hypothetical protein	SACOL1162	S.aureus . COL
Unknown function	13	General	sensor histidine kinase family protein	SACOL0202	S.aureus . COL
Unknown function	13	General	guanylate kinase	SACOL1220	S.aureus . COL
Unknown function	13	General	anti protein (phenol soluble modulin)	SACOL1187	S.aureus . COL
Unknown function	13	General	anti protein (phenol soluble modulin)	SACOL1186	S.aureus . COL

Table 5.2. Dwon regulated genes is mastitis *S. aureus* 764 internalized into mammary cells at normal O_2 condition

Main Role	Sub-role	Product	Locus	Org. Descr.
Cell envelope	Other		SAB0392	S.aureusRF122
Cell envelope	Other		SAB2220c	S.aureusRF122
Cell envelope	Other		SAB1660c	S.aureusRF122
Cell envelope	Other		SACOL0193	S.aureus COL
Cell envelope	Other		SAB0396	S.aureusRF122
Cell envelope	Other		SAB2529c	S.aureusRF122
Cell envelope	Other		SACOL2443	S.aureus COL
Cell envelope	Biosynth &degradn of surface polysac & lipopolysac		SAB0901	S.aureusRF122
Hypothetical proteins	Domain		SACOL0445	S.aureus COL
Energy metabolism	Glycolysis/gluconeogenesis	ldh2	SACOL2618	S.aureus COL
Energy metabolism	Glycolysis/gluconeogenesis	ldh2	SACOL2618	S.aureus COL
Energy metabolism	Fermentation		SACOL0215	S.aureus COL
Energy metabolism	Anaerobic	ldh2	SACOL2618	S.aureus COL
Energy metabolism	Anaerobic	ldh2	SACOL2618	S.aureus COL
Energy metabolism	Electron transport	NADH dehase(*nuoF*	SACOL0494	S.aureus COL
Cellular processes	Detoxification		SACOL2641	S.aureus COL
Cellular processes	Detoxification		SACOL0451	S.aureus COL
Cellular processes	Cell division		SACOL1202	S.aureus COL
Cellular processes	Toxin production and resistance		SACOL2465	S.aureus COL
Regulatory functions	Other		SAB1289c	S.aureusRF122
Regulatory functions	Other		SACOL2585	S.aureus COL
Regulatory functions	Protein interactions		SACOL1939	S.aureus COL
Fatty acd & phospholipd meta.	Degradation		SACOL0214	S.aureus COL
Fatty acd & phospholipd meta.	Degradation		SACOL0213	S.aureus COL
Fatty acd & phospholipd meta.	Biosynthesis		SACOL2482	S.aureus COL
Fatty acd & phospholipd meta.	Prophage functions		SAB1727c	S.aureusRF122
Fatty acd & phospholipd meta.	Transposon functions		SAB1167c	S.aureusRF122
Fatty acd & phospholipd meta.	Other		SACOL2465	S.aureus COL
Unknown function	General		SAB1660c	S.aureusRF122
Unknown function	General		SACOL1941	S.aureus COL
Unknown function	General		SACOL2484	S.aureus COL
Signal transduction	PTS		SAB0782	S.aureusRF122
Signal transduction	PTS		SACOL2552	S.aureus COL
Biosy cofactrs, pros grps& carirs	Glutathione and analogs		SACOL2641	S.aureus COL
Biosy cofactrs, pros grps& carirs	Other		SACOL2579	S.aureus COL
Transcription	DNA-dependent RNA polymerase		SAB2012c	S.aureusRF122
Transcription	DNA-dependent RNA polymerase		SACOL1222	S.aureus COL
DNA metabolism	DNA replication, recombination, and repair		SACOL1150	S.aureus COL
Amino acid biosynthesis	Glutamate family		SACOL0514	S.aureus COL
Central intermediary metabolis	Other		SACOL2577	S.aureus COL
Protein fate	Degradation of proteins, peptides, & glycopeptides		SAB2566	S.aureusRF122

Genome-wide expresson profiles of mastitis *S. aureus* isolate 764 internalized into mammary cells at reduced oxygen condition

One of the most upregulated genes under this condition were those of pathogenesis subset of the cellular processes class; particularly, toxin production and resistance including alph and delta toxins, leukocidins, aerolysins, as well as superantigen like toxins, as shown in Figure 5.3 and Table 5.3. The rest included a few cation-induced transport processes such as sodium-, iron-, and ammonium-mediated transport, as well as some binding surface proteins such as fibrinogen binding related proteins, iron-regulated surface determinant (Isd) protein genes *isdA*, *isdC, isdH*, and serine aspirate dipeptide repeat protein genes (*sdrH*). In addition, unlike that of strain Wright under the same condition, some electron transport and tricarboxylic acid cycle components were expressed. Chromosomally encoded genes such as phenol soluble modulins were also expressed. Finally, a large class of proteins with unknown function (14%), and some hypothetical were also expressed.

The largest gene class that was down regulated under reduced oxygen comprised mostly metabolic and biosynthetic pathways including pyrimidine, purine, and large number of cell division proteins that were arrested. However, the major difference between the activities of this strain in this condition compared to that at normal oxygen, was the down regulation of components of glycolytic or fermentative pathways such as GntR, as well as those of alternative sugar metabolism as indicated by the inactivity of fructose and sucrose metabolism. Thus, the unique property under this condition was the general metabolic inactivity (Table 5.4). In addition, hypothetical proteins (21.4%) and a significant percent of genes with unknown functions (11.9%) (Figure 5.4) were also down regulated.

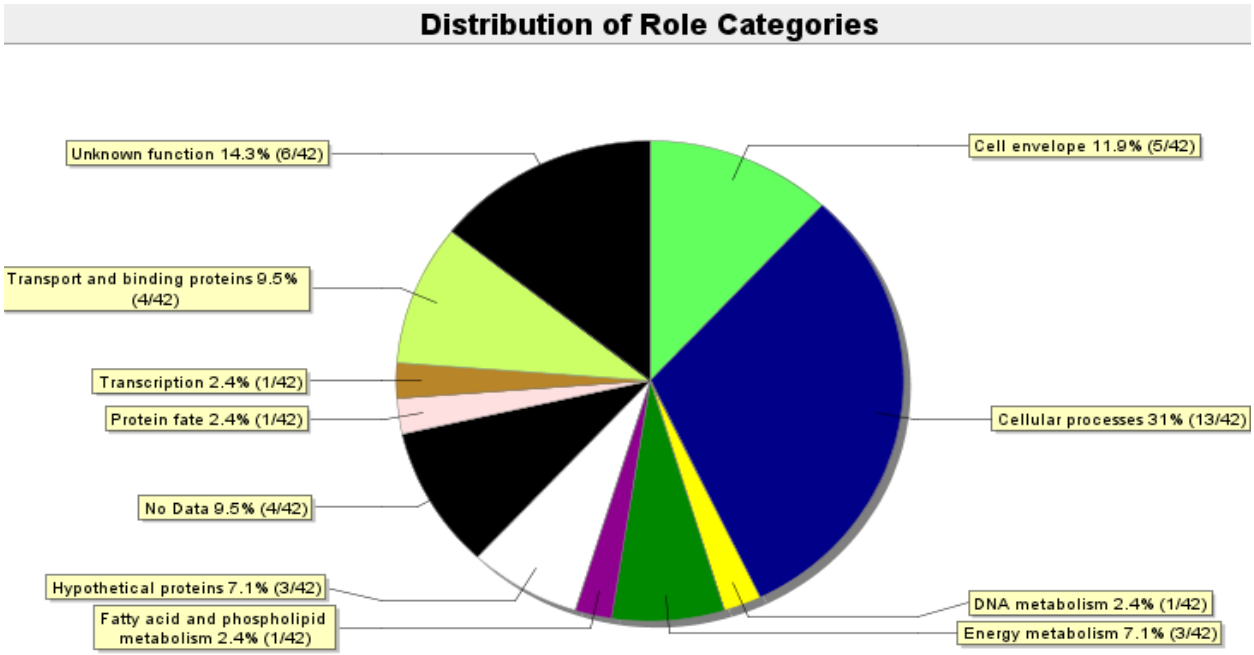


Figure 5.3 Subsets of upregulated gene classes in mastitis *S. aureus* 764 internalzied into mammary cells at reduced O_2 condition

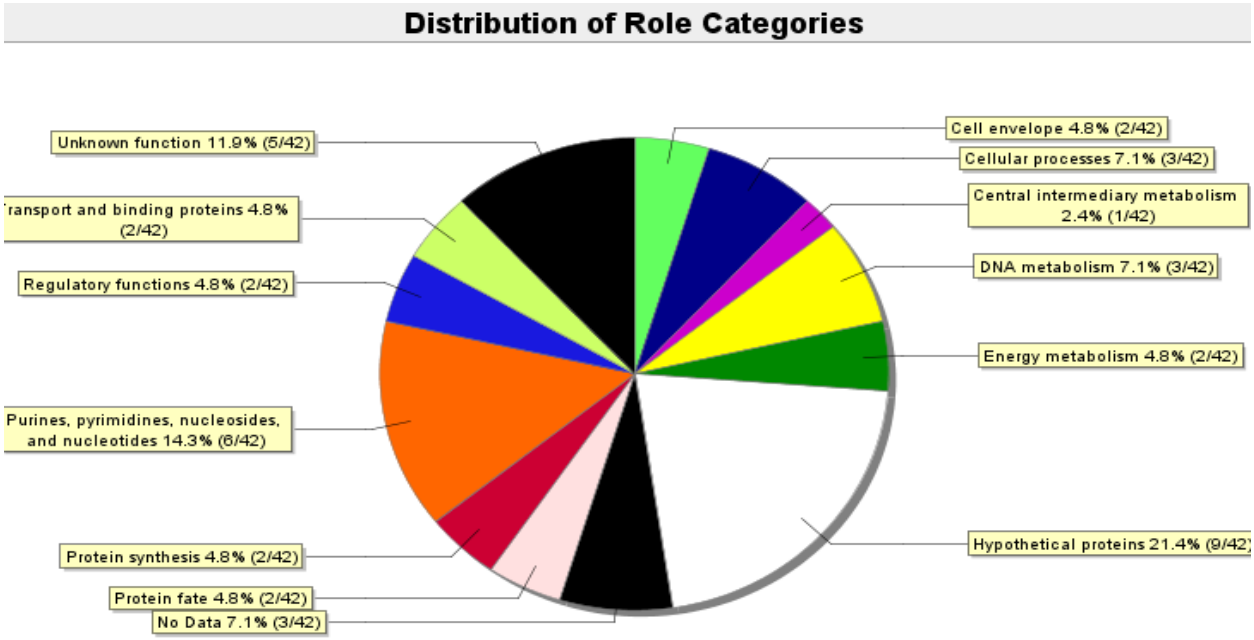


Figure 5.4. Subsets of downregulated gene classes in mastitis *S. aureus* 764 intenalized into mammary cells at reduced O_2 condition.

5.3 Upregulated genes in mastitis *S. aureus* 764 internalized into mammary cells at reduced O_2 condition

Main Role	MRC	Sub-role	SRC	RID	Product	Locus	Org. Descr.
Cellular processes	13	Pathogenesis	11	187	hly, alpha-hemolysin	SACOL1173	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	hly, alpha-hemolysin	SACOL1173	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	superantigen-like pt	SACOL1180	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	superantigen-like pt	SACOL1180	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	cys ptease precu SspB	SACOL1970	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	Leukocidin	SACOL2004	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	Leukocidin	SACOL2004	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	Aeroly/leuko	SACOL2006	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	Aeroly/leuko	SACOL2006	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	hld, delta-hemo	SACOL2022	S. aureus COL
Cellular processes	13	Pathogenesis	11	187	hld, delta-hemo	SACOL2022	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	Hypothetical	SACOL1152	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	hly, alpha-hemoly	SACOL1173	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	hly, alpha-hemolysin	SACOL1173	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	superantigen-like pt	SACOL1180	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	superantigen-like pt	SACOL1180	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	Leukocidin	SACOL2004	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	Leukocidin	SACOL2004	S. aureus COL
Cellular processes	13	Toxin pd& resist	12	94	Aeroly/leuko	SACOL2006	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	Aeroly/leuko	SACOL2006	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	TcmP, putative	SACOL2009	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	hld, delta-hemo	SACOL2022	S. aureus COL
Cellular processes	13	Toxin prd& resist	12	94	hld, delta-hemo	SACOL2022	S. aureus COL
Unknown function	6	Enzy unkno speci	2	703	exonuclease	SACOL1954	S. aureus COL
Unknown function	6	Enzy unkno speci	2	703	hydrolase	SACOL2021	S. aureus COL
Unknown function	6	General	4	157	phenol-soluble modul	SACOL1186	S. aureus COL
Unknown function	6	General	4	157	phenol-soluble modul	SACOL1187	S. aureus COL
Unknown function	6	General	4	157	fgen bindg-related pt	SACOL1220	S. aureus COL
Cell envelope	5	Other	5	88	isdA	SACOL1140	S. aureus COL
Cell envelope	5	Other	5	88	isdC	SACOL1141	S. aureus COL
Cell envelope	5	Other	5	88	iron-binding prot	SACOL2010	S. aureus COL
Cell envelope	5	Other	5	88	SdrH, putative	SACOL2019	S. aureus COL
Trans bindg prot	4	Cations&F- copds	4	145	iron-binding prot	SACOL1144	S. aureus COL
Trans bindg prot	4	Cations&F- copds	4	145	iron-binding prot	SACOL2010	S. aureus COL
Trans bindg prot	4	Cations&F- copds	4	145	Na-transport protein	SACOL2011	S. aureus COL
Trans bindg prot	4	Cations&F- copds	4	145	Ammonium trans prot	SACOL2031	S. aureus COL
Energ metabolism	3	Electron transport	2	112	sdhC,succ dehase,TCA	SACOL1158	S. aureus COL
Energymetabolism	3	Electron transport	2	112	sdhC,succ dehase,TCA	SACOL1158	S. aureus COL
Energymetabolism	3	TCA cycle	2	120	sdhC,succ dehase,TCA	SACOL1158	S. aureus COL
Energ metabolism	3	TCA cycle	2	120	sdhC,succ dehase,TCA	SACOL1158	S. aureus COL
Energymetabolism	3	Ami acids amines	1	109	arcC1,carbamate kinase	SACOL1182	S. aureus COL
Hypo proteins	3	Conserved	3	156	ZapA, cell divi protein	SACOL1151	S. aureus COL
Protein fate	1	Deg pts, pept, glycopep	1	138	cys ptease precu SspB	SACOL1970	S. aureus COL
F acid, p-lipid met	1	Degradation	1	177	hlb,phospholipase C	SACOL2003	S. aureus COL

5.4 Downregulated genes in mastitis *S. aureus* 764 internalized into mammary cells at reduced O_2 condition

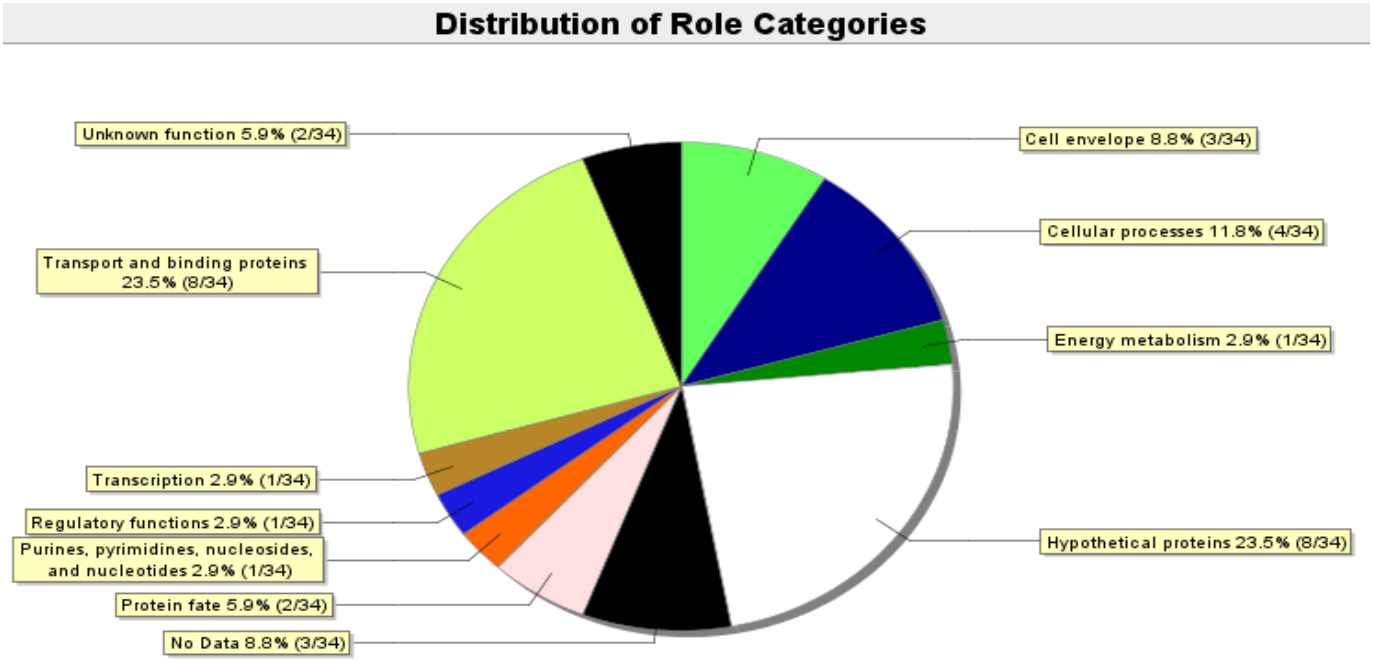
Main Role	MRC	Sub-role	SRC	Product	Locus	Org. Descr.
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1190	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1218	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1226	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1968	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1985	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1987	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1991	S aureus COL
Hypothetical proteins	9	Conserved	9	hypothetical	SACOL1992	S aureus COL
Hypothetical proteins	9	Conserved	9	hypoth,envir sens	SACOL1993	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	pyrB,pyr, aa, meta	SACOL1212	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	pyrC,pyr, aa, meta	SACOL1213	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	carA, pyr, aa, meta	SACOL1214	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	carB,pyr, aa, meta	SACOL1215	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	pyrF,pyr, meta	SACOL1216	S aureus COL
Pur, pyr, nucleo-sides &tides	6	Pyr biosyn	6	pyrE,pyr meta	SACOL1217	S aureus COL
Unknown function	5	General	3	fib-bind related prot	SACOL1164	S aureus COL
Unknown function	5	General	3	MraZ, cell divi prot	SACOL1191	S aureus COL
Cellular processes	3	Cell division	3	cell division prot	SACOL1197	S aureus COL
Cellular processes	3	Cell division	3	ftsA, cell divis prot	SACOL1198	S aureus COL
Cellular processes	3	Cell division	3	ftsZ, cell divis prot	SACOL1199	S aureus COL
DNA metabolism	3	DNA repl, recomb,&repair	3	FtsQ, cell divis prot	SACOL1150	S aureus COL
DNA metabolism	3	DNA repl, recomb,&repair	3	MutS2, DNA repair	SACOL1154	S aureus COL
DNA metabolism	3	DNA repl, recomb,&repair	3	ligA, DNAligase, NAD dept	SACOL1965	S aureus COL
Protein fate	2	Protein folding& stabilzn	2	groEL, chaperonin	SACOL2016	S aureus COL
Protein fate	2	Protein folding& stabilzn	2	groES,cochaperonin	SACOL2017	S aureus COL
Energy metabolism	2	Sugars	1	fructokinase,suc,fru met	SACOL2028	S aureus COL
Energy metabolism	2	Biosyn &degrad polysacch	1	cscA, sucr6p hylase	SACOL2029	S aureus COL
Cell envelope	2	Biosyn &degrad polysacch	2	mraY,Peptidog syn	SACOL1195	S aureus COL
Cell envelope	2	Biosyn &degrad polysacch	2	murD,Peptidog syn	SACOL1196	S aureus COL
Transport & binding proteins	2	aa, peptides and amines	1	proline permease	SACOL1963	S aureus COL
Transport & binding proteins	2	Nucleosides, pur &pyr	1	uraA,uracilpermase	SACOL1211	S aureus COL
Regulatory functions	2	DNA interactions	1	GntR,(ccpA,PMP inactive)	SACOL1997	S aureus COL
Regulatory functions	2	RNA interactions	1	pyrR	SACOL1210	S aureus COL
Central metabolism	1	Phosphorus compounds	1	ppaC,inorg pyro-pase	SACOL1982	S aureus COL

Expression profiles of human *S. aureus* Wright internalized into mammary cells at normal oxygen condition

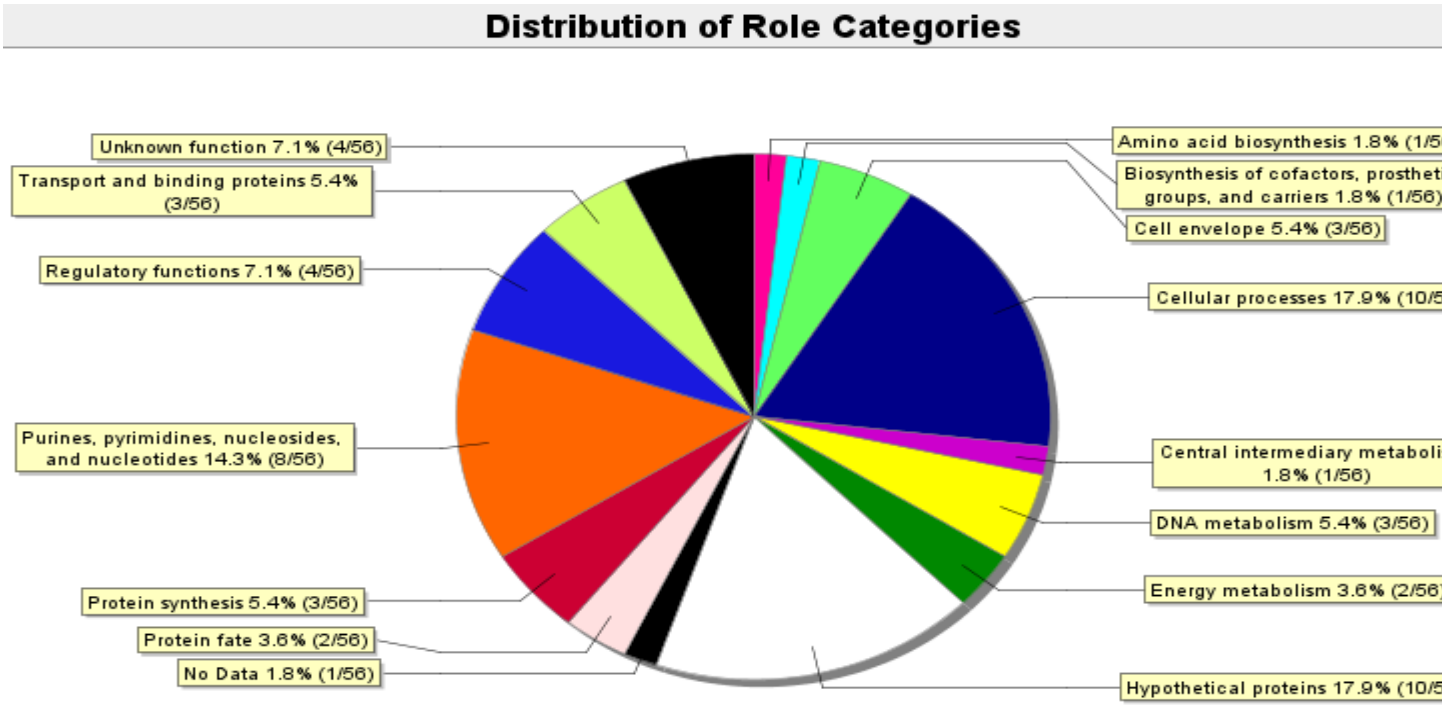
During the invasion of mammary cells under normal oxygen levels, gene subsets belonging to transport and binding and cellular processes showed the most expression levels of all classes in strain Wright (Figure 5.5, Table 5.5). However, in pathogenesis subset of the cellular processes genes, i.e., production of toxins and exoproteins were minimally expressed. Most of the transport processes involved sodium-dependent dicarboxylate transport (*sdcS*), iron-dependent transport, as well as iron-regulated surface determinant proteins *isdA*, *isdC*. In addition, some cysteine protease precursors, aerolysins, agr, and toxins were also expressed. However, metabolic and biosynthetic genes were minimally expressed; and among the least expressed was the tricarboxylic acid cycle enzyme, *sdhB*, succinate dehydrogenase. Similar to other experiments, the majority of genes under this condition belonged to the hypothetical class, and some with unknown functions.

In the subsets of down regulated genes, there was general limitations in expression in all classes including biosynthesis and central metabolic inactivity, glucose and alternative sugar metabolism, cell division proteins genes (*FtsQ*), and chromosomally encoded genes such as phenol-soluble modulins (Figure 5.6, Table 5.6). The most common down regulations were the those belonging to the toxin production including exfoliative toxin, aerolysin leukocidin, and delta hemolysins. This coincided with dwon regulation in the in *agr* expression levels. Compared to the up regulated gene classes, the majority of genes in the cellular processes class which contains pathogenesis were down regulated (confirming that active pathogenesis was not produced). Also, as expected the transport and binding proteins were shown as minimal; similarly, cell adhesion, environmental sensing, and intracellular matrix binding such as fg-binding-related protein. Most of the genes that were not expressed were hypothetical proteins.

.



5.5 Subsets of upregulated gene classes in human *S. aureus* Wright internalized into mammary cells at normal oxygen condition



5.6 Subsets of downregulated gene classes in human *S. aureus* Wright internalized into mammary cells at normal oxygen condition

5.5 Upregulated genes in human *S.aureus* Wright internalized into mammary cells at normal O2 condition

Main Role	Sub-role	SRC	Product	Locus
Hyp proteins	Conserved	8	redox-sensing transcrip repress,REX	SACOL2035
Trans bindg prot	aa, pept amines	1	proline permease	SACOL1963
Trans bindg prot	Unknown substrate	4	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Unknown substrate	4	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Unknown substrate	4	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Unknown substrate	4	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Cations F-compds	7	iron-binding prot	SACOL1144
Trans bindg prot	Cations F-compds	7	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Cations F-compds	7	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Cations F-compds	7	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Cations F-compds	7	Na-dept dicooHate transport sdcS	SACOL1979
Trans bindg prot	Cations F-compds	7	iron-binding prot	SACOL2010
Trans bindg prot	Cations F-compds	7	Na-transport protein	SACOL2011
Cellular processes	Pathogenesis	4	cystei ptease precur SspB, Staphopain A, deg elastin, up by agr	SACOL1970
Cellular processes	Pathogenesis	4	Aeroly/leuko	SACOL2006
Cellular processes	Pathogenesis	4	Aeroly/leuko	SACOL2006
Cellular processes	Pathogenesis	4	agrA	SACOL2026
Cellular processes	Toxin prod resist	2	Aeroly/leuko	SACOL2006
Cellular processes	Toxin prod resist	2	Aeroly/leuko	SACOL2006
Cell envelope	Other	3	isdA	SACOL1140
Cell envelope	Other	3	isdC	SACOL1141
Cell envelope	Other	3	iron-binding prot	SACOL2010
Unknown function	General	2	protein phosphatase	SACOL1231
Unknown function	General	2	pseudogene	SACOL2015
Protein fate	Degn prots, peptid,& glycopeptides	2	cys ptease precu SspB	SACOL1970
Protein fate	Degn of prots, peptid,& glycopeptides	2	pseudogene	SACOL2007
Energy metabolism	TCA cycle	1	sdhB, succinate dehydrogenase	SACOL1160
Pur, pyr, nucleosides & tides	nucleosides & tides interconversions	1	gmk, guanylate kinase	SACOL1221
Transcription	DNA-dependent RNA polymerase	1		SACOL1222
Regulatory functions	Other	1	agrA	SACOL2026

5.6 Downregulated genes in human *S. aureus* Wright internalized into mammary cells at normal O2 condition

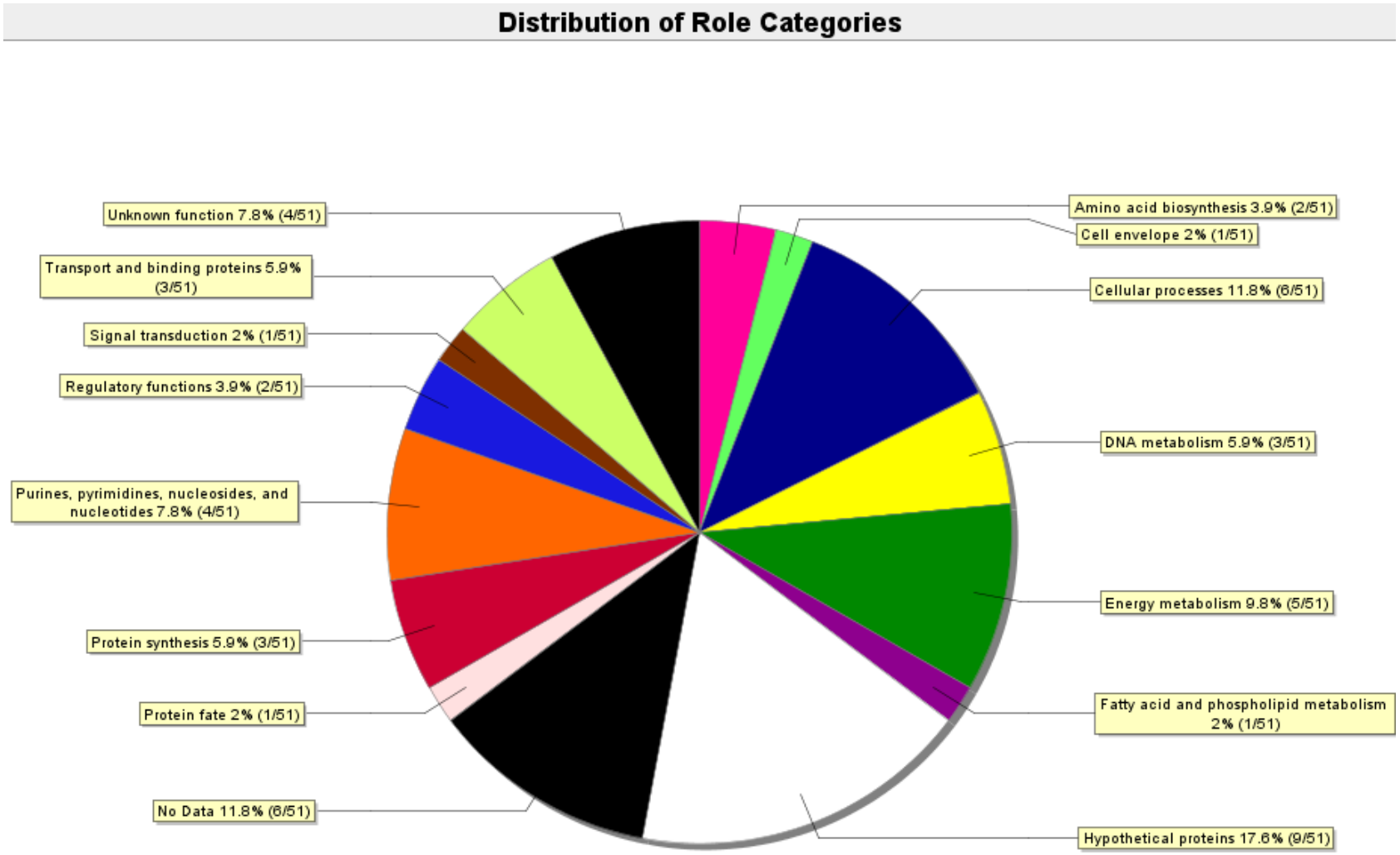
Main Role	Sub-role	Product
Cellular processes	Cell adhesion	camS
Cellular processes	Cell adhesion	camS
Cellular processes	Pathogenesis	exfoliative toxin, puta
Cellular processes	Pathogenesis	exfoliative toxin, puta
Cellular processes	Pathogenesis	Aeroly/leuko
Cellular processes	Pathogenesis	Aeroly/leuko
Cellular processes	Pathogenesis	hld, delta-hemo
Cellular processes	Pathogenesis	hld, delta-hemo
Cellular processes	Pathogenesis	AgrC
Cellular processes	Other	camS
Cellular processes	Other	camS
Cellular processes	Toxin production and resistance	exfoliative toxin, puta
Cellular processes	Toxin production and resistance	exfoliative toxin, puta
Cellular processes	Toxin production and resistance	exfoliative toxin, puta
Cellular processes	Toxin production and resistance	exfoliative toxin, puta
Cellular processes	Toxin production and resistance	hld, delta-hemo
Cellular processes	Toxin production and resistance	hld, delta-hemo
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	pyrB,pyr, aa, meta
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	pyrC,pyr, aa, meta
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	carA, pyr, aa, meta
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	carB,pyr, aa, meta
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	pyrF,pyr, meta
Pur, pyr, nucleo-sides & tides	Pyrimidine ribonucleotide biosyn	pyrE,pyr meta
Pur, pyr, nucleo-sides & tides	Purine ribonucleotide biosynthesis	purB,adenylosuccinate lyase
Pur, pyr, nucleo-sides & tides	Purine ribonucleotide biosynthesis	purB,adenylosuccinate lyase
Unknown function	General	fg-binding-related protein
Unknown function	General	phenol-soluble modul
Unknown function	General	phenol-soluble modul
Regulatory functions	DNA interactions	scrR,sucrose operon repressor
Regulatory functions	DNA interactions	scrR,sucrose operon repressor
Regulatory functions	RNA interactions	pyrR
Regulatory functions	Other	AgrC
Regulatory functions	Small molecule interactions	scrR,sucrose operon repressor
Regulatory functions	Small molecule interactions	scrR,sucrose operon repressor
DNA metabolism	DNA repln, recombn, & repair	FtsQ, cell divis prot
DNA metabolism	DNA repln, recombn, & repair	ligA, DNAligase, NAD dept
Transport and binding proteins	Unknown substrate	ABC transporter
Transport and binding proteins	Unknown substrate	ABC transporter
Transport and binding proteins	Nucleosides, pur and pyri	uracil permease
Cell envelope	Biosyn degradn of mureins &peptido	mraY,Peptidog syn
Cell envelope	Biosyn degradn of mureins &peptido	ABC transporter
Cell envelope	Biosyn degradn of mureins &peptido	Mur ligase family protein
Energy metabolism	Amino acids and amines	arcB1,Arg and proline meta
Energy metabolism	Amino acids and amines	arcC1,carbamate kinase
Central intermediary metabolism	Other	NO synthase, O2ase subunit
Amino acid biosynthesis	Glutamate family	arcB1,Arg and proline meta
Biosy cofactrs, pros gps,&cariers	Heme, porphyrin, and cobalamin	Hypothetical

Expression profiles of human *S. aureus* Wright internalized into mammary cells at reduced oxygen condition

At this condition (Figure 5.7), the highest up regulated expression rate was shown by the group of genes responsible for the cellular processes (subsets of genes for chemotaxis, cell adhesion, and pathogenesis Table 5.7). Similar to strain 764 at normal oxygen with 5% CO2, this strain (Wright) also showed up regulated detoxification genes para-nitrobenzyl esterase pnbA but also increased expression of alkyl hydroperoxide reductase subunit C (*ahpC*) under this condition. Expression of Glycolysis (Fruc-1-6,biphosphate aldolase, *fdaB*), anaerobic genes (formate acetyltransferase genes *pflA* SACOL0205 and *pflB* SACOL0204), pyruvate formate-lyase-activating enzyme *pflA*, PTS system components IIBC, and chromosomal genes phenol soluble modulins (SACOL1187), were also seen, but with much lower rates compared to that of strain 764. However, for the regulatory functions, in addition to MerR family transcriptional regulator (SACOL2517), an RF122 bovine strain specific enterotoxin N (SAB1697c) was produced. Only a limited toxin production (*hld,* delta hemolysin locus SACOL2022, superantigen like SACOL0468) and cell envelope and transport and binding protein showed increased expression. Higher proportion of genes in the conserved subsets were hypothetical proteins except for redox-sensing transcriptional repressor Rex (SACOL2035), thioredoxin (*trxA*, SACOL1155), and a bovine pathogenicity island protein (SAB0351). No much multiplication, metabolic activities, or cell division genes were expressed.

The subsets of genes that showed down regulation in expression patterns under reduced oxygen condition are shown in Figure 5.8, and the details of main role and sub roles of each gene class are shown in Table 5.8. Notably, the number of hypothetical proteins that were down regulated was significantly lower than that reported in the up regulated portion under reduced oxygen condition, for this important class. No data was obtainable from 6 and 10 loci in the analysis of their up and down regulation, respectively. In Figure 5.8, the cellular processes subset was the most unique in terms of their transcriptional profiles; many important genes were differentially transcribed. One of the most interesting transcriptional profiles found under reduced oxygen, was the universally down regulated activity of the *sarU* in all biological roles. In Table 5.8, the

activity of this regulator was shown under pathogenesis (*sarU*, SACOL2507), toxin production and resistance (with similar down regulation in superantigen-like protein, SACOL0473, SACOL0472, SACOL0469), in cell envelope, in hypothetical proteins, as well as in the regulatory functions where highest decrease in expression was observed (see *sarU* locus Table 5.8). Similarly, the GntR (locus SACOL0518) family of transcriptional regulator, which belongs to regulatory functions, was also down regulated by reduced oxygen. In the electron transport class, NADH dehydrogenase subunit 5 (*nuoF,* SACOL0494) was down regulated, while in the central intermediary metabolism D-lactate dehydrogenase (SACOL2535) was not active. Furthermore, one of the genes responsible for "adaption to atypical condition" called a large-conductance mechanosensitive channel, (*mscL*, locus SAB1205c) which is a stress response gene was not expressed. Finally, among the down regulated genes, many hypothetical proteins were bovine strain specific



Fiugre 5.7 Subsets of up regulated gene classes in human *S. aureus* Wright internalized into mammary cells at reduced oxygen condition

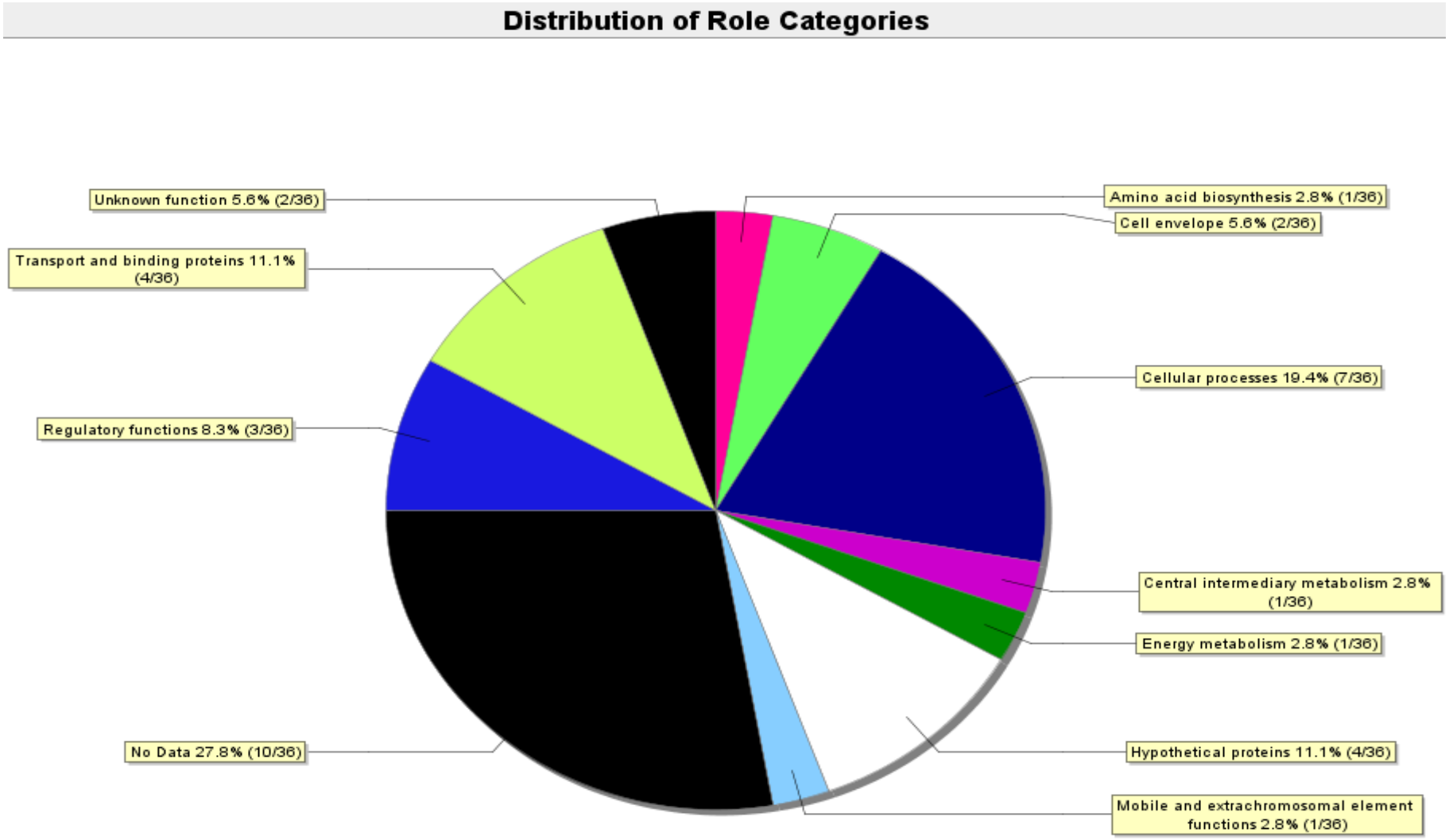


Figure 5.8. Subsets of down regulated gene classes in human *S. aureus* Wright internalized into mammary cells at reduced oxygen condition

Table 5.7. Upregulated genes of human *S. aureus* Wright internalized into mammary cells at reduced O_2 condition

Main Role	Sub-role	Product	Locus	Org. Descr.
Hypothetical proteins	Domain		SACOL0445	S aureus COL
Cellular processes	Detoxification	ahpC, alkyl hydperoxide reductase C	SACOL0452	S aureus COL
Cellular processes	Detoxification	pnbA, para-nitrobenzyl esterase	SACOL2459	S aureus COL
Cellular processes	Pathogenesis	hld	SACOL2022	S aureus COL
Cellular processes	Pathogenesis	hld	SACOL2022	S aureus COL
Cellular processes	Cell division	ftsZ, cel div pt	SACOL1199	S aureus COL
Cellular processes	Toxin production and resistance	superantigen like	SACOL0468	S aureus COL
Cellular processes	Toxin production and resistance	hld	SACOL2022	S aureus COL
Cellular processes	Toxin production and resistance	hld	SACOL2022	S aureus COL
Energy metabolism	Glycolysis/gluconeogenesis	*fdaB* fru-1,6-bp aldase	SACOL2622	S aureus COL
Energy metabolism	Sugars		SACOL2515	S aureus COL
Energy metabolism	Anaerobic	pflB, fomaT acetyltranfas	SACOL0204	S aureus COL
Energy metabolism	Anaerobic	pflA,pyr for-lyase-actg enz	SACOL0205	S aureus COL
Energy metabolism	Electron transport	trxA, thioredoxin	SACOL1155	S aureus COL
Pur, pyri, nucleosid, and nucleotid	Pyrimid ribonucleot biosynthesis	carB,pyr met	SACOL1215	S aureus COL
Pur, pyri, nucleosid, and nucleotid	Pyrimid ribonucleot biosynthesis	guaB, pur met	SACOL0460	S aureus COL
Pur, pyri, nucleosid, and nucleotid	2'-Deoxyribonucleotide meta	nrdG, ano2 prot	SACOL2634	S aureus COL
Pur, pyri, nucleosid, and nucleotid	2'-Deoxyribonucleotide meta	DNA met	SACOL2635	S aureus COL
Unknown function	General		SACOL0191	S aureus COL
Unknown function	General		SACOL0507	S aureus COL

Unknown function	General	modulins	SACOL1187	S aureus COL
Unknown function	General		SACOL0021	S aureus COL
DNA metabolism	DNA repli, recombi, &repair		SACOL0509	S aureus COL
DNA metabolism	DNA repli, recombi, &repair		SACOL0016	S aureus COL
DNA metabolism	DNA repli, recombi, &repair		SACOL0438	S aureus COL
Protein synthesis	Ribosomal prots: synth &modifn		SACOL0437	S aureus COL
Protein synthesis	tRNA aminoacylation		SACOL1961	S aureus COL
Protein synthesis	tRNA aminoacylation		SACOL1960	S aureus COL
Transport and binding proteins	Carbos, organic alcohs,& acids		SACOL0516	S aureus COL
Transport and binding proteins	Unknown substrate		SAB2408	S.aureus RF122
Transport and binding proteins	Cations & iron carrying compds		SACOL1952	S aureus COL
Amino acid biosynthesis	Glutamate family		SACOL0514	S aureus COL
Amino acid biosynthesis	Glutamate family		SACOL0515	S aureus COL
Regulatory functions	DNA interactions	MerR Reg	SACOL2517	S aureus COL
Regulatory functions	Other	Entrotox N	SAB1697c	S.aureus RF122
Protein fate	Protein modification and repair		SACOL0205	S aureus COL
Fatty acid &phospholip meta	Degradation		SACOL2459	S aureus COL
Cell envelope	Other	hemolysin	SAB2051c	S.aureus RF122
Signal transduction	PTS		SACOL0516	S aureus COL

Table 5.8. Down regulated genes in human *S. aureus* Wright internalized into mammary cells at reduced O_2 condition

Main Role	Sub-role	Input Acc.	Product	Locus	Org. Descr.
Cellular processes	Adaptations to atypical conditions	SAB1205c		SAB1205c	S.aureus RF122
Cellular processes	Cell division	SACOL1197		SACOL1197	S. aureus COL
Cellular processes	Pathogenesis	SACOL2507		SACOL2507	S. aureus COL
Cellular processes	Pathogenesis	SACOL2507		SACOL2507	S. aureus COL
Cellular processes	Toxin production and resistance	SACOL0473		SACOL0473	S. aureus COL
Cellular processes	Toxin production and resistance	SACOL0472		SACOL0472	S. aureus COL
Cellular processes	Toxin production and resistance	SACOL2507		SACOL2507	S. aureus COL
Cellular processes	Toxin production and resistance	SACOL2507		SACOL2507	S. aureus COL
Cellular processes	Toxin production and resistance	SACOL0469		SACOL0469	S. aureus COL
Hypothetical proteins	Conserved	SAB0141		SAB0141	S.aureus RF122
Hypothetical proteins	Conserved	NWMN_2472		NWMN_2472	S.aureusNM
Hypothetical proteins	Conserved	SAB0239		SAB0239	S.aureus RF122
Hypothetical proteins	Conserved	SAB1720c		SAB1720c	S.aureus RF122
Transport and binding proteins	Amino acids, peptides and amines	SACOL0010		SACOL0010	S. aureus COL
Transport and binding proteins	Amino acids, peptides and amines	SACOL2441		SACOL2441	S. aureus COL
Transport and binding proteins	Unknown substrate	SAB0138		SAB0138	S.aureus RF122
Transport and binding proteins	Unknown substrate	SACOL0506		SACOL0506	S. aureus COL
Regulatory functions	DNA interactions	SACOL2507		SACOL2507	S. aureus COL
Regulatory functions	DNA interactions	SACOL2507		SACOL2507	S. aureus COL
Regulatory functions	DNA interactions	SACOL0518		SACOL0518	S. aureus COL
Regulatory functions	Oth	SACOL2507		SACOL2507	S. aureus COL
Regulatory functions	Other	SACOL2507		SACOL2507	S. aureus COL
Cell envelope	Other	SACOL2505		SACOL2505	S. aureus COL
Cell envelope	Other	SACOL0482		SACOL0482	S. aureus COL
Unknown function	Enzymes of unknown specificity	SACOL0510		SACOL0510	S. aureus COL
Unknown function	Enzymes of unknown specificity	SACOL0531		SACOL0531	S. aureus COL
Amino acid biosynthesis	Glutamate family	SACOL0169		SACOL0169	S. aureus COL
Energy metabolism	Electron transport	SACOL0494	NADH dehase(*nuoF*	SACOL0494	S. aureus COL
Central intermed metabolism	Other	SACOL2535		SACOL2535	S. aureus COL
Mobile& extrachro elem functions	Transposon functions	SACOL0464		SACOL0464	S. aureus COL

DISCUSSION

In recent years, *S. aureus* has been the focus of intensive research efforts because of its importance as a major human and animal pathogen. Intensive genome sequencing of different *S aureus* strains would potentially reveal how different lineages adapt to different host species and evolve into subpopulations. In this comprehensive study, we have shown for the first time, to the best of our knowledge, the high-throughput whole-genome expression profiles of *S. aureus* directly on clinical isolates of mastitis- and human-specific *S. aureus* during invasion into mammary epithelial cells, that lead to the identification of subsets of genes expressed. SarU mediated agr activation, intensive environmental sensing pathways, as well as numerous detoxification and coupling of fermentative metabolism to virulence have been observed with mastitis isolate under normal oxygen condition. On the other hand, metabolic inactivity indicative of persistency was more evident under reduced oxygen in both strains. The isolates examined represented major clones as identified by comparative analysis using *clfA* and the gold standards PFGE and spa-typing (Said et al., 2010). Thus, we have provided new perspectives on staphylococcus pathogenesis and host-specialization.

During invasion of bovine mammary cells incubated under normal oxygen condition, specific subsets of genes were expressed in mastitis isolate 764 (Figure 5.1). The main and sub role categories are shown in Table 5.1. The two component regulatory systems, *agr* and *sarU*, were the most expressed genes. Interestingly, this increased expression of the *agr* and its activating gene, *sarU*, suggested that induction of *agr* was through this environmentally sensitive DNA binding protein of the *sarA* family, indicating a potential alternative pathway in the mammary gland (Yarwood et al., 2007; Bronner et al., 2004). This showed that the intracellular microenvironment of the mammary cells, and not the bacterial density, was responsible for the alternative pathways leading to activation of original *agr* signal that resulted in the elevated production of toxins and repression of protein A (*spa*) (Cheung et al., 2008); perhaps through the proposed biphasic intracellular induction of virulence (Shompole et al., 2003). This is consistent with the fact that Spa does not form immune complex with cow's serum (Atkins et al., 2008), and therefore might be inefficient in mastitis situation.

Simultaneous transcription of several regulatory and some metabolic pathways suggested co-ordinated expression of virulence factors. Expression of the sensor kinase VraS implied cell protection from lysis because the genes regulated by this system are induced upon cell-wall inhibitory effects (Kuroda et al., 2003; Belcheva and Golemi-Kotra, 2008). These observations, along with the expression of *pnbA* (para-nitrobenzyl esterase) for hydrolysis of beta-lactam antibiotic esters and *betA* and *betB* for dehydrogenation of choline and betaine, respectively, imply that detoxification and adaptation to host intramammary conditions has occurred. Expression of the TetR family of transcriptional regulators supports that mastitis strain was well adapted to dairy cows consistent with finding that these regulators are abundant in animal pathogens, and were suggested to be involved in adaptation to complex and changing environments (Ramos et al., 2005). Moreover, expression of autolytic cysteine proteases *sspB* is consistent with survival strategy at the intracellular phase of invasion, as explained below (McGavin et al., 1997; Nickerson *et al.*, 2007; 2008). These intracellular survival strategies are further supported by the expression of *yycG* and *yycI* loci. The (YycG/YycF, renamed as WaIK/WalR) two-component system has been conserved in Gram-positive bacteria, including *S. aureus* and is essential for coupling cell viability and wall metabolism, and is therefore, an attractive targets for novel classes of antimicrobial compounds (Dubrac et al., 2007; Dubrac and Msadek, 2004). WalK is one of the only two proteins, the other being PhoR, that contain a PAS (PER-ARNT-SIM) domain which is involved in sensing signals such as oxygen, light, redox potential, or the presence of specific ligands (Taylor and Zhulin, 1999). In addition, the WalKR system is shown to control virulence genes such as *sdrD*, (a sialoproreln-bindingprotein), and *ebpS*, (an elastin-bindingprotein) (Dubrac and Msadek, 2004).

Coupling of metabolic pathways to virulence gene expression might also have occurred. Energy metabolism was induced through glycolysis, fermentation, and anaerobic pathways, as the genes encoding the enzymes of these pathways were expressed such as *ldh* (Garrard and Lascelles, 1968). This is in agreement with Seidl et al.,'s (2009) finding that in the presence of glucose, genes required for glycolysis were up regulated, while those of tricarboxylic acid cycle and amino acid utilization were repressed by the catabolite control protein A (CcpA). The *ccpA* gene also controls

virulence gene expression and its activation in this study was evident by the up regulation of the pentose monophosphate pathway (PMP) operon (GntR) which is regulated by the *ccpA* (Seidl et al.,'s (2009). It appears that glycolysis was a preferred metabolic pathway even under aerobic condition since the oxidative form of glycolysis, the PMP pathway, was activated generating NADPH and synthesis of phosphorylated 5 carbon sugar; perhaps for components of RNA and DNA, as well as ATP, NADH, FAD, and coenzyme A. Thus, up regulation of a key enzyme in mixed acid fermentation, pyruvate formate lyase (PFL), in addition to similar hypoxic metabolic pathways suggests a direct link between these highly preferred fermentative metabolism and virulence in *S. aureus*. This has been recently demonstrated in *Streptococcus pneumonia* (Yesilkaya et al., 2009). The CO2 used for cell culture incubator could have induced this route of metabolism. However, much of the energy gained from glycolysis seemed to be used for sensing, protection, and pathogenesis because most of the biosynthetic processes were down. The GntR operon controls the multidrug resistance pumps, like NorB, which facilitates bacterial microenvironmental fitness in abscesses (Ding et al., 2008). Another differentially expressed core chromosomal genes were the phenol soluble modulin. A recent comprehensive study by Li et al., (2009) concluded that differential expression of phenol soluble modulins and alpha toxins, and not the acquisition of virulent factors by horizontal transfer were responsible for the increased virulence in the recently emerged clones. In addition, the large majority of genes with unknown or hypothetical functions were expressed as shown in the class of hypothetic proteins. Future annotation projects would certainly gain more insights into the regulatory process of *S. aureus* in bovine mastitis.

Several lines of evidence suggested intracellular transcriptional control of gene expression through environmental sensing. For example, expression of the MerR suggested activation of sigma factor, in response to environmental stimuli, such as oxidative stress, metal ions or antibiotics (Brown et al., 2003). Similarly, a significant increase in *araC* expression, which is part of the commonly conserved *araC/xylS* protein family in gram positive bacteria, indicates simultaneous coupling of carbon metabolism, stress response, and pathogenesis (Gallegos et al., 1997). However, a recent study suggested that although conserved regulators may be present in different genera or

species, their actual regulatory activity could be only lineage specific (Janga and Pérez-Rueda, 2009). Thus, mastitis isolate 764 showed reduced metabolic activity on the account of invasion or persistence, adaptation, and environmental sensing pathways.

The down regulated genes in strain 764 under normal oxygen condition were biosnythesis, cell envelope, and transport and binding protein. Mastitis strain RF122 specific genes in classes of cell envelope and transport and binding proteins were also reduced (Table 5.2). Interestingly, RF122-specific fibronectin-binding protein (locus SAB1289c) was found truncated, inconsistent with the suggestion that it promoted colonization (Brouillette et al., 2004). Considering the host- and tissue-specialization of the species, this might be due to the above authors' use of derivatives of prototype laboratory strains (*S. aureus* 8325) in a mouse model of mastitis, or perhaps due to their correlation of the CFU (colony forming units) from mammary homogenates to colonization, when multiplication is best related to the invasive phase where binding proteins are prone to degradation or blocked by proteases (McGavin et al., 1997). Furthermore, the propionate metabolic pathways and tricaroxylic acid cycle genes were down regulated by this isolate under normal oxygen condition.

The up regulated response of mastitis isolate 764 under reduced oxygen was similar to that its expression profile at normal oxygen in terms of toxin production and pathogenesis (Figure 5.3, Table 5.3). However, unlike under normal oxygen, surface binding proteins such as fibrinogen binding related proteins, iron-regulated surface determinant genes (Isd) proteins *isdA*, *isdC*, *isdH*, and serine aspirate dipeptide repeat proteins (*sdrH*) were expressed. These proteins were found expressed only under iron deplete (Vasi et al., 2009; Skaar et al., 2004) consistent with the free-iron restricted bovine mammary gland intracellular condition, and were found to play important roles during intracellular survival in agreement with the findings of Vaudaux et al. (2002), and Sendi et al. (2009). This along with the significant down regulation in metabolic and biosynthetic pathways (Figure 5.4, Table 5.4) is very consistent with an intracellular behavior of the bovine isolate 764. Annotation of the large number of hypothetical (21%) and those with unknown functions (11.9%) may gain more insights.

Similar to the mastitis strain, strain Wright internalized at normal oxygen levels also exhibited typical *in vivo* profile of gene expression (Figure 5.5, Table 5.5). Expression of the *isd*

operon, the cysteine protease precursors, and aerolysins, indicated invasive intracellular phase of the infection. It is well known that *agr* induces several secreted proteins that promote the invasion and intracellular spread, including serine protease, lipase, fibrinolysin, alphahemolysin, beta-hemolysin, delta-hemolysin, enterotoxin B, and toxic shock syndrome toxin since all were found reduced by *agr* mutations (Bjkorklind and Arvidson, 1980). In this respect, protease activity inhibits binding such as fibronectin binding phenotype (MacGavin et al., 1997) as a result of activation of the metalloprotease aureolysin leading to activation of the serine protease operon *ssp*ABC (Nickerson *et al.*, 2007; 2008). Staphopain A (ScpA) and Staphopain B (SspB) are expressed in respective operons *scp*AB and *ssp*ABC, where the Staphopain genes *scp*A and *ssp*B are followed by genes encoding Staphostatins ScpB and SspC, which inhibit the respective Staphopains during binding stages of growth (Rzychon *et al.*, 2003). Furthermore, expression of sodium-dependent dicarboxylate transport system (*sdcS*) and tricarboxylic acid intermediate genes (Figure 5.5, Table 5.5) showed that aerobic respiration had occurred, unlike that of mastitis strain under the same condition, and that in both strains most of the energy was directed to pathogenesis. These are supported by down regulations in biosynthesis, cell adhesion, intracellular matrix binding such as fibrinogen-binding-related protein (Figure 5.6, Table 5.6).

The human associated strain Wright expressed a few fermentative pathway genes including *fdaB, pflA,B*, and thioredoxin at reduced oxygen (Figure 5.7, Table 5.7). However, a limited toxin production was observed hinting for a limitation in *agr* activity. Nevertheless, the activity of environmentally-sensitive alternative transcriptional regulators had occurred. An additional detoxification product, the oxidative stress resistance enzyme alkyl hydroperoxide reductase (*ahpC*) was expressed (Table 5.5). This is an alternative hydrogen peroxide scavenger indicating catalase activity was reduced under this condition, and that was compensated by *ahpC* gene product. Expression of the PerR regulon genes, alkyl hydroperoxide reductase (*ahpC)* and catalase (*katA*) implied reduced oxygen and persistence properties (Cosgrove et al., 2007). In addition, similar expression patterns have been shown to be induced due to osmotic pressure and NaCL (Armstrong-Buisseret et al., 1995), and acid shock (Bore et al., 2007). Finally, an alternative pathway of environmental sensing involving activation of *sigB* through *merR*

could have resulted (Brown et al., 2003). Even though these stress inducing conditions are consistent with the favored fermentative pathways, it is not clear why would *S. aureus* prefer them.

Experssion profiles of strain Wright genes down regulated under reduced oxygen are shown in Figure 5.8 and Table 5.8. One of the most interesting pattern found was the universally down regulated *sarU* activity in all biological roles. Specifically, this occurred in sub roles (Table 5.8) toxin production and resistance, with substantial down regulation in superantigen-like protein, (SACOL0473, SACOL0472, SACOL0469), in cell envelope, in hypothetical proteins, as well as in the regulatory functions where highest decrease in expression was observed. Similarly, cessation of fermentative pathway with concomitant down regulation in major stress response gene responsible for "adaption to atypical condition" called a large-conductance mechanosensitive channel, (*mscL*, locus SAB1205c) (Kouwen et al., 2009), has indicated a major shift in gene expression. These include down regulation of NADH dehydrogenase which would strongly suggest down regulation of *walKR* system (Dubrac et al., 2007; Dubrac and Msadek, 2004). More importantly, down regulation of *ccpA* was evidenced by down regulation of GntR which is regulated by the *ccpA* (Seidl et al., 2009). It follows that members of the GntR, AraC/XylS and GalR/LacI families that generally respond to environmental changes that affect the carbohydrate metabolism were also arrested (Perez-Rueda and Collado-Vides, 2000). These observations suggest that reduced oxygen potentially promoted persistence properties evidenced by metabolic inactivity (Shompole et al., 2003; Wesson et al., 1998).

In summary, invasion of human and mastitis associated *S. aureus* into mammary epithelial cells under normal and reduced oxygen conditions showed subsets of genes specific to the condition as well as strain. Irrespective of the condition the mastitis strain showed expression profiling for binding, energy production, and *agr*-related toxin production, expoproteins, as well as the regulator *agr* and *sarU* genes, indicating colonization, adaptation, and intracellular spreading phenotypes. However, while increased expression of genes belonging to adaptation, detoxification, and environmental sensing were observed under normal oxygen, up regulation of iron- and fibrinogen-binding genes and down regulation of metabolism was evident under reduced oxygen. On the other hand, the human strain showed significant reduction in metabolic and toxin

genes compared to mastitis strain; particularly, at reduced oxygen condition indicating persistence profiles. Thus, we have provided high-throughput tarnscriptome profiles of mammary epithelial cell internalized mastitis- and human-associated *S. aureus* strains and revealed potential regulatory pathways in this study. In addition, role category-based analysis has identified specific subsets of genes that were coordinately transcribed. Future annotation project of the large number of hypothetical proteins and those with unknown functions would potentially reveal a more precise network of regulatory subsets that are involved in the mammary gland microenvironment enabling significant insights into the basic mechanism of host- and tissue-specialization as well as in the selection of suitable therapeutic candidates.

ACKNOWLEDGEMENTS

This study was supported by funding from a discovery grant from Natural Science and Engineering Research Council of Canada and from the Canadian Bovine Mastitis Research Network in a collaborative research with PFGRC of the J. Craig Venter Institute, Maryland USA.

REFERENCES

Armstrong-Buisseret L, Cole MB, Stewart GS. 1995. A homologue to the *Escherichia coli* alkyl hydroperoxide reductase AhpC is induced by osmotic upshock in *Staphylococcus aureus.* Microbiology 141:1655-1661.

Atkins K, Burman J, Chamberlain E, Cooper J, Poutrel B, Bagby S, Jenkins A, Feil E, van den Elsen J. 2008. *S. aureus* IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. Mol Immunol 45: 1600-1611.

Bayles K, Wesson C, Liou L, Fox L, Bohach G, Trumble W. 1998. Intracellular *Staphylococcus aureus* escapes the endosome and induces apoptosis in epithelial cells. Infect Immun 66: 336-342.

Belcheva A, Golemi-Kotra D. 2008. A close-up view of the VraSR two-component system. A mediator of *Staphylococcus aureus* response to cell wall damage. J Biol Chem 283:12354-12364.

Bjkorklind, A., S. Arvidson**.** 1980. Mutants of *Staphylococcus aureus* affected in the regulation of exoprotein synthesis. FEMS Microbiol Lett **7:**203–206.

Bore E, Langsrud S, Langsrud O, Rode TM, Holck A. 2007. Acid-shock responses in *Staphylococcus aureus* investigated by global gene expression analysis. Microbiology 153:2289-2303.

Bronner S, Monteil H, Prévost G. 2004. Regulation of virulence determinants in *Staphylococcus aureus*: complexity and applications. FEMS Microbiol Rev 28:183-200.

Brouillette E, Talbot BG, Malouin F. 2003. The Fibronectin-Binding Proteins of *Staphylococcus aureus* May Promote Mammary Gland Colonization in a Lactating Mouse Model of Mastitis. Infect Immun 71:2292–2295.

Brown NL, Stoyanov JV, Kidd SP, Hobman JL. 2003. The MerR family of transcriptional regulators. FEMS Microbiol Rev 27:145-163.

Bubeck-Wardenburg J, Palazzolo-Ballance A, Otto M, Schneewind O, DeLeo F. 2008. Panton-Valentine leukocidin is not a virulence determinant in murine models of community-associated methicillin-resistant *Staphylococcus aureus* disease. J Infect Dis 198: 1166-1170.

Cheung A, Nishina K, Trotonda M, Tamber S. 2008. The SarA protein family of *Staphylococcus aureus*. Int J Biochem Cell Biol 40: 355-361.

Cosgrove K, Coutts G, Jonsson IM, Tarkowski A, Kokai-Kun JF, Mond JJ, Foster SJ. 2007. Catalase (KatA) and alkyl hydroperoxide reductase (AhpC) have compensatory roles in peroxide stress resistance and are required for survival, persistence, and nasal colonization in *Staphylococcus aureus*. J Bacteriol. 189:1025-1035.

Ding Y, Onodera Y, Lee JC, Hooper DC. 2008. NorB, an efflux pump in *Staphylococcus aureus* strain MW2, contributes to bacterial fitness in abscesses. J Bacteriol 190 21:7123-9.

Dubrac S, Msadek T. 2004. Identification of genes controlled by the essential YycGlYycF two-component system of *Staphylococcus aureus*. J Bacteriol 186:1175-1181.

Dubrac S, Boneca IG, Poupel O Msadek T. 2007. New insights into the WaIK/WalR (YycGlYycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. J Bacteriol 189:8257-8269.

Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. Arac/XylS family of transcriptional regulators. Microbiol Mol Biol Rev 61:393-410.

Garzoni C, Francois P, Huyghe A, Couzinet S, Tapparel C, Charbonnier Y, Renzoni A, Lucchini S, Lew DP, Vaudaux P, Kelley WL, Schrenzel J. 2007. A global view of *Staphylococcus aureus* whole genome expression upon internalization in human epithelial cells. BMC Genomics. 8:171.

Garrard W, J Lascelles. 1968. Regulation of *Staphylococcus aureus* lactate dehydrogenase. J Bacteriol 95:152–156.

Grundmeier M, Hussain M, Becker P, Heilmann C, Peters G, Sinha B. 2004. Truncation of fibronectin-binding proteins in *Staphylococcus aureus* strain Newman leads to deficient adherence and host cell invasion due to loss of the cell wall anchor function. Infect Immun 72: 7155-7163.

Huynh HT, Robitaille G, Turner JD. 1991. Establishment of bovine mammary epithelial cells (MAC-T): an *in vivo* model for bovine lactation. Exp Cell Res 197:191-199.

Janga SC, Perez-Rueda E. 2009. Plasticity of transcriptional machinery in bacteria is increased by the repertoire of regulatory families. Comput Biol Chem 33:261-268.

Kouwen TR, Trip EN, Denham EL, Sibbald MJ, Dubois JY, van Dijl JM. 2009. The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168. Antimicrob Agents Chemother 53:4702-4711

Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K.2003. Two-component system VraSR positively modulates the regulation of cell-wall biosynthesis pathway in *Staphylococcus aureus*. Mol Microbiol. 49:807-821.

Li M, Diep B, Villaruz A, Braughton K, Jiang X, DeLeo F, Chambers H, Lu Y, Otto M. 2009. Evolution of virulence in epidemic community-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci U S A 106: 5883-5888.

McAleese F, Walsh E, Sieprawska M, Potempa J, Foster T. 2001. Loss of clumping factor B fibrnogen binding activity by *Staphylococcus aureus* involves cessation of

transcription, shedding and cleavage by metalloprotease. J Biol Chem 276: 29969-29978.

McGavin MJ, Zahradka C, Rice K Scott JE. 1997. Modification of the Staphylococcus aureus fibronectin binding phenotype by V8 protease. Infect Immun 65: 2621-2628.

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

Nickerson NN, Joag V, McGavin MJ. 2008. Rapid autocatalytic activation of the M4 metalloprotease aureolysin is controlled by a conserved N-terminal fungalysin-thermolysin-propeptide domain. Mol Microbiol 69: 1530–1543.

Nickerson NN, Prasad L, Jacob L, Delbaere LT, McGavin MJ. 2007. Activation of the SspA serine protease zymogen of *Staphylococcus aureus* proceeds through unique variations of a trypsinogen-like mechanism and is dependent on both autocatalytic and metalloprotease mtalloprotease-specific processing. J Biol Chem 282: 34129–34138.

Novick, RP. 2003. Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol 48:1429–1449.

Perez-Rueda E, Collado-Vides J. 2000. The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. Nucleic Acids Research 28: 1838-1847.

Qazi SN, Harrison SE, Self T, Williams P, Hill PJ. 2004. Real-time monitoring of intracellular *Staphylococcus aureus* replication. J Bacteriol 186:1065–1077.

Ramos J, Martı´nez-Bueno M, Molina-Henares AJ, TeraW, Watanabe K, Zhang X, Gallegos MT, Brennan R, and Tobe R. 2005. The TetR family of transcriptional repressors. Microbiol Mol Biol Rev 69:326-56.

Rossney A, Shore A, Morgan P, Fitzgibbon M, O'Connell B, Coleman D. 2007. The emergence and importation of diverse genotypes of methicillin-resistant *Staphylococcus aureus* (MRSA) harboring the Panton-Valentine leukocidin gene (pvl) reveal that pvl is a poor marker for community-acquired MRSA strains in Ireland. J Clin Microbiol 45: 2554-2563.

Rzychon M, Sabat A, Kosowska K, Potempa J, Dubin A. 2003. Staphostatins: an expanding new group of proteinase inhibitors with a unique specificity for the regulation of

Staphopains, *Staphylococcus spp*. cysteine proteinases. Mol Microbiol 49**:** 1051–1066.

Said KB, Ismail J, Campbell J, Mulvey MR, Bourgault A-M, Messier S, Zhao X. 2010. Regional profiling for determination of genotype diversity of mastitis-specific *Staphylococcus aureus* lineage in Canada by use of clumping Factor A, Pulsed-Field Gel Electrophoresis, and *spa* typing. J Clin Microbiol 48:375-386.

Seidl K, Müller S, François P, Kriebitzsch C, Schrenzel J, Engelmann S, Bischoff M, Berger-Bächi B. 2009. Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus***.** BMC Microbiol. 18: 9:95.

Sendi P, Proctor RA. 2009. *Staphylococcus aureus* as an intracellular pathogen: the role of small colony variants. Trends Microbiol 17:54-88.

Senn M, Bischoff M, von Eiff C, Berger-Bächi B. 2005. sigmaB activity in a *Staphylococcus aureus* hemB mutant. J Bacteriol 187: 7397-7406.

Shompole S, Kim TH, Linda E L, Katarzyna D, Gregory AB and Kenneth WB. 2003. Biphasic intracellular expression of *Staphylococcus aureus* virulence factors and evidence for Agr-mediated diffusion sensing. Mol. Microbiol. 49:919–927.

Skaar EP, Schneewind O. 2004. Iron-regulated surface determinants (Isd) of *Staphylococcus aureus*: stealing iron from heme. Microbes and Infection 6:390-397.

Somerville G, Beres S, Fitzgerald J, DeLeo F, Cole R, Hoff J, Musser J. 2002. *In vitro* serial passage of *Staphylococcus aureus*: changes in physiology, virulence factor production, and *agr* nucleotide sequence. J Bacteriol 184: 1430-1437.

Ster C, Gilbert F, Cochard T, Poutrel B. 2005. Transcriptional profiles of regulatory and virulence factors of *Staphylococcus aureus* of bovine origin: oxygen impact and strain-to-strain variations. Mol Cell Probes 19: 227-235.

Taylor BL, Zhulin IB. 1999. PAS domains: internal sensors of oxygen, redox potential and light. Microbiol Mol Biol Rev 63:479-506.

Tuchscherr L, Buzzola F, Alvarez L, Caccuri R, Lee J, Sordelli D. 2005. Capsule-negative *Staphylococcus aureus* induces chronic experimental mastitis in mice. Infect Immun 73: 7932-7937.

Vaudaux P, Francois P, Bisognano C, Kelley WL, Lew DP, Schrenzel J, Proctor RA, McNamara PJ, Peters G, Von Eiff Ch. 2002. Increased expression of clumping factor

and fibronectin-binding proteins by *hemB* mutants of *Staphylococcus aureus* expressing small colony variant phenotypes. Infect. Immun 70:5428–5437.

Visai L, Yanagisawa N, Josefsson E, Tarkowsk A, Pezzali I, Rooijakkers SHM. Foster TJ, Speziale P. 2009. Immune evasion by *Staphylococcus aureus* conferred by iron-regulated surface determinant protein IsdH. Microbiology 155:667-679.

Wertheim H, Walsh E, Choudhurry R, Melles D, Boelens H, Miajlovic H, Verbrugh H, Foster T, van Belkum A. 2008. Key role for clumping factor B in *Staphylococcus aureus* nasal colonization of humans. PLoS Med 5: e17.

Wesson C, Liou L, Todd K, Bohach G, Trumble W, Bayles K. 1998. *Staphylococcus aureus* Agr and Sar global regulators influence internalization and induction of apoptosis. Infect Immun 66: 5238-5243.

Yarwood JM, Paquette KM, Tikh IB, Volper EM, Greenberg EP. 2007. Generation of virulence factor variants in *Staphylococcus aureus* biofilms. J Bacteriol 189:7961-7967.

Yesilkaya H, Spissu F, Carvalho SM, Terra VS, Homer KA, Benisty R, Porat N, Neves AR, Andrew PW. 2009. Pyruvate formate lyase is required for pneumococcal fermentative metabolism and virulence. Infect Immun 77:5418-27.

Zhang K, McClure J, Elsayed S, Tan J, Conly J. 2008. Coexistence of Panton-Valentine leukocidin-positive and -negative community-associated methicillin-resistant *Staphylococcus aureus* USA400 sibling strains in a large Canadian health-care region. J Infect Dis 197: 195-204.

CHAPTER VI. GENERAL DISCUSSION AND CONCLUSION

Staphylococcus aureus has been one of the most common causes of bovine intramammary infections (IMI) worldwide. Similarly, *S. aureus* is an important pathogen causing human infections. However, the factors and mechanisms leading to the emergence of separate lineages for each host and organ, has been quite complex. The primary infection control strategy starts with specific identification and typing to differentiate infectious strains from commensals and transient colonizers. For *S. aureus*, this has been a difficult task mainly because the species possesses a highly conserved genome yet specific strains rapidly transform into hypervirulent clones. For instance, in the year 2005, a specific lineage group caused more invasive diseases than the combined rates of diseases caused by bacterial species with most transformable genomes; namely, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Neisseria meningitidis*, and *Haemophilus influenza* and *S. aureus* mortality rates well exceeded that of the HIV (Bancroft, 2007, Klevens et al., 2007). Recent realization of this and the global economic impact have alerted an international effort recommending objective strain typing as a pre-requisite for basic research in all branches of microbiology.

Selection of molecular typing approaches depends upon the goal defined beforehand. Methods used for investigating short term outbreaks differ from those employed in long-term surveillance programs; hence, combination of methods has been recommended (Criso´stomo et al., 2001; Villari et al., 1998; Sabria-Leal et al., 1994). However; for use in multidisciplinary approaches, markers that influence genetic traits such as conserved adhesin coding loci with highly variable domains have been preferred. These markers can successfully couple straightforward epidemiological investigations with phenotypes and in this way typing can generate diagnostically useful markers or those associated with infectious disease pathogenesis or colonisation potential.

It is critical in *S. aureus* research to understand the rapidly changing epidemiology and the basic mechanisms underlying emergence of specific types with superior virulence. This is because strain types produced by major methods such as multilocus sequence typing (MLST) might not provide any information on clinical or phenotypic profiles. The design of MLST is based on neutrally evolving essential genes encoding

central metabolic enzymes (Maiden et al., 1998) and as such, predicts only the population structure revealing common ancestoral lines through the evolutionary history. A relevant example was revealed by the comparative analysis of two complete genomes of *S. aureus* strains MW2 and MSSA476 which were isolated from the USA and UK, respectively. Although both strains have identical MLST (sequence type ST1) the clinical and pathogenic phenotypes were quite different; the former caused a fatal paediatric bacteraemia, while the latter was isolated from a case of osteomyelitis, and the patient was fully recovered (Holden et al., 2004; Baba et al., 2002). However, these two strains were identical by MLST. The variable and accessory genes were responsible for the variations in pathogenic properties that caused different clinical phenotypes. Grouping by these genes have been highly preferred, as mentioned above, in case of *S. aureus.* Consequently, three different categories of genomic regions defined as the core genome, core variable genome, and accessory genome have been proposed as a result of multi-genome comparative analysis (Lindsay et al.; 2006). The core-variable and accessory-variable genomes encode surface proteins that mediate adaptation and constitute good candidates as molecular typing markers.

In this study, we first chose *clfA* gene as a potential molecular typing marker. The suitability of the repeat domain in the *clfA* locus as the molecular typing marker was examined using three reference strains, 19 mastitis isolates, and 24 human isolates. These isolates and strains were vigorously passaged in different laboratory media and conditions as well as for invasion experiments into mammary epithelial cells, testing variations in repeat copy numbers of each isolate by PCR and sequencing. Stability of the repeats under rigorous treatments coupled to their inherent variability among different strains from different host species proved *clfA* suitable as a typing marker. Interestingly, the patterns of R domains polymorphisms including all published sequences of human and bovine *S. aureus* available at the time (Figures 2.3, 2.4A, 2.4B) followed host-specificity; and therefore, the suitability of the *clfA* locus for coupling host-specific typing and strain phenotypes was achieved. We have also found, for the first time, that human isolates were highly polymorphic, while those of mastitis were homogenous. This implied an organ-specificity in the *clfA* locus was further examined in subsequent studies. The reasonable stability of the *clfA* repeats and its tendency to acquire a defined length ranges in invasive

strains (dictated by the number of copies in the R domain as the rest is conserved), as well as the unique clonality of mastitis isolates prompted to investigate organ-specificity of the locus.

In chapter III, to examine the organ-specific repeat variations in the *clfA*, we have analyzed a significant number of isolates (95), from different organs in human patients and from bovine mastitis, for the number of copies as well as the nucleotide sequences in their hypervariable R domain using PCR, sequencing, multiple sequence alignment, and phylogenetic relationships. A very highly significant correlation was obtained between the length of the *clfA* (number of repeat copies) and the host organ from where isolates were isolated. As shown in Table 3.2, twenty out of the 23 sputum isolates had lower copy-numbers of 43-48, while 21 out of the 24 skin isolates had 55-63 copies. In addition, repeat-types and overall sequence pattern-groups were highly consistent, and together provided a grouping pattern similar to that obtained by major molecular methods such as MLST. An additional advantage of *clfA* method was the phenotype-specificity coupled to strain differentiation. For instance, sequence alignments (Figure 3.2) and phylogenetic analysis (Figure 3.3) grouped host- and tissue-specific isolates in respective clusters. More importantly, the mastitis isolates tested from two different countries were highly clonal for the *clfA* with 76% showing a single copy number (52 copies of repeats). These findings clearly indicate selection of certain *clfA* length to certain organs. Since *clfA* is a universal surface adhesin important in virulence, these specificities have to do with the colonization and invasiveness, evolution, and spreading strategies of *S. aureus*.

In the above studies we have identified and evaluated *clfA* for host- and organ-specific-grouping marker that couples colonization and invasion phenotypes. There is accumulated evidence to support our findings. ClfA belongs to the MSCRAMMs responsible for adaptations (Foster and Hook, 1998) and encoded by the species-specific core-variable genomic domain (*Ori environ*) that is responsible for staphylococcal-speciation (Takeuchi et al., 2005; Lindsay et al., 2006; Dordet-Frisoni et al., 2007). Many reports cite the uniquely abundant ClfA activity during infection (Wolz et al., 1996, 2002; Josefsson et al., 1998), and that it mediates virulence even in absence of its ligand fibrinogen (Palmqvist et al., 2004a). Similarly, under bovine-mammary mimicked low oxygen condition, enhanced *clfA* and *spa* expression was detected (Ster et al., 2005). This

has been recently substantiated in the intramammary microenvironment using anti-ClfA antibody, where significant ClfA activity was found even in persistent forms of the cocci (Tuchscherr et al., 2008). Furthermore, this protein has antiphagocytic properties comparable to that of Spa (protein A), and that could occur independent of its ligand, fibrinogen (Higgins et al., 2006). More specifically ClfA has been shown to be responsible for the intracellular density and spreading of the cocci (Ahmed et al., 2001), and as such, has been one of genetic requirements in the abscess formation and persistence (Palmqvist et al., 2004b; Cheng et al., 2009).

After the development, evaluation, and testing of *clfA* for host- and organ-specificity, we proceeded to standardize this approach by examining its performance and discriminatory power in another comprehensive comparative analysis experiment with the gold standard methods, the pulsed field gel electrophoresis (PFGE) and *spa* typing. As indicated in chapter IV, for this project, a total of 93 *S. aureus* were examined, 87 of which were isolates from mastitis cases representing Western provinces, Ontario, Quebec, and Eastern provinces of Canada with a final goal of identifying dominant repeat types with common properties for use in whole-genome expression profiling. Interestingly, clfA performed quite well achieving an excellent discriminatory power with an index of discrimination of 0.9, identical to that of the PFGE. All of the tested isolates belonged to two major PFGE groups A (48.3%) and D (43.7) (Figure 4.1 and Table 4.1). These groups mainly represented the Eastern and Western Canada strain types, respectively. Consistent with earlier findings that only a limited number of clones cause mastitis (Kapur et al., 1995; Fitzgerald et al., 1997) we provide, for the first time, a precise identification of the dominant types as well as their regional distributions in Canada.

More importantly, the dominant PFGE types were further distinguished into sub types by the *clfA*. Four dominant *clfA*-types X, Q, C, and Z were obtained, which comprised 82% and 43% of PFGE groups A and D, respectively, and varied only within a narrow range of copy number of 46 to 52 copies, implying clonal-selection probably due to local selection pressures (Table 4.1). This is in agreements with other's findings that *S. aureus* evolves into subclonal populations in different geographic locations and that these can be detected by repeat based typing methods (Saei et al., 2009; Su et al., 1999). Sequence alignments and phylogenetic analysis (Figures 4.2A and 4.3) further confirmed

the groupings based on copy number and were highly consistent with earlier findings as determined by major molecular methods (Baba et al., 2008; Highlander et al., 2007). There was agreement between the three methods in overall strain clustering; however, *clfA* provided a further finer level of subtyping dominant types with an excellent performance and discriminatory power. Distinct mastitis-associated *S. aureus* strain types with common properties representing the dominant clones were characterized for use in whole-wide genome transcripome profiling.

The two dominant clones identified through a comprehensive sequence of investigations reported in chapters II, III, and IV were finally used to measure the changes in gene expression after their internalization into mammary epithelial cells. These clones were segregated based on their genotype that was essentially coupled to their potential colonization and adherence properties inherent in the *clfA*. As shown in Figure 5.1 and Table 5.1 for the mastitis strain 764 under normal oxygen condition, *agr*-activation had potentially ocurred through the *sarU* pathway (possible routes are indicate in Figure 1.9) (Yarwood et al., 2007; Bronner et al., 2004) resulting in the activation of original *agr* signal leading to elevated production of toxins and repression of *spa* (Cheung et al., 2008). This along with the inability of Spa to form immune complex with cow's serum (Atkins et al., 2008) and the antiphagocytic activity of the ClfA that matches that of Spa (Higgins et al., 2006), suggested compensation of protein A activity by the clumping factor in hoofed animals. Expression of genes responsible for cell wall protection, detoxification, cell viability, stress response, as well as two component systems that simultaneously couple environmental sensing and virulence gene expression suggested a well established intracellular response (Taylor and Zhulin, 1999; Belcheva and Golemi-Kotra, 2008; Dubrac and Msadek, 2004). Similarly, expression of the MerR and sigma factor, suggested increased responses to environmental stimuli (Brown et al., 2003).

One of the unique expression profiles in mastitis 764 under normal oxygen was the preferential coupling of glycolytic pathways with virulence properties despite adequate oxygen level. Consistent with Seidl et al.,'s (2009) finding, catabolite control protein (CcpA) was activated and the genes for fermentative metabolism and virulence such as the pentose monophosphate pathway (PMP) operon (GntR) including the NorB which facilitates bacterial fitness in abscesses (where oxygen content is usually limited),

were all up regulated (Ding et al., 2008). Another example was the overexpression of *araC* for simultaneous coupling of carbon metabolism, stress response, and pathogenesis (Gallegos et al., 1997). The subsets of down regulated genes in mastitis isolate 764 (as indicated in Figure 5.2 and Table 5.2) included the propionate metabolic pathways and the tricaroxylic acid cycle genes. Interestingly, a link between anaerobic metabolism and virulence has been recently established in *Streptococcus pneumonia* (Yesilkaya et al., 2009). However, these changes in expression patterns seemed quite consistent with an *in vivo* response for optimal performance at the mildly acidic infection site (Weinrick et al, 2004), although quite different sets of regulators were involved in this invasion study.

The up regulated gene classes in mastitis isolate 764 under reduced oxygen were similar to its profiles at normal oxygen in toxin production and pathogenesis (Figure 5.3, Table 5.3). Nevertheless, surface binding proteins such as fibrinogen binding related proteins, iron-regulated surface determinant genes (Isd) proteins *isdA*, *isdC*, *isdH*, and serine aspirate dipeptide repeat proteins (*sdrH*) were expressed at reduced oxygen. Up regulation of these genes, known to be expressed only under iron deplete, (Vasi et al., 2009; Skaar et al., 2004) is consistent with the free-iron restricted bovine mammary gland intracellular condition. They have also been found to play important roles during intracellular survival in agreement with the findings of Vaudaux et al., (2002), and Sendi et al., (2009). This along with the significant down regulation in metabolic and biosynthetic pathways under reduced oxygen (Figure 5.4, Table 5.4) is very consistent with an intracellular behavior of the bovine isolate 764. Annotation of the large number of hypothetical (21%) and those with unknown functions (11.9%) may gain more insights.

Human associated strain Wright internalized at normal oxygen levels also exhibited typical *in vivo* profiles of gene expression (Figure 5.5, Table 5.5). Expressions of the *isd* operon, the cysteine protease precursors, and aerolysins, were up regulated. Protease activity has been found to inhibit binding such as fibronectin binding phenotype (MacGavin et al., 1997) as a result of activation of the metalloprotease aureolysin leading to activation of the serine protease operon *ssp*ABC (Nickerson *et al.*, 2007; 2008). Staphopain A (ScpA) and Staphopain B (SspB) are expressed in respective operons *scp*AB and *ssp*ABC, where the Staphopain genes *scp*A and *ssp*B are followed by genes

encoding Staphostatins ScpB and SspC, which inhibit the respective Staphopains during binding stages of growth (Rzychon *et al.*, 2003). Furthermore, expression of sodium-dependent dicarboxylate transport system (*sdcS*) and tricarboxylic acid intermediate genes (Figure 5.5, Table 5.5) showed that aerobic respiration had occurred, unlike that of mastitis strain under the same condition, and that in both strains most of the energy was directed to pathogenesis. These are supported by down regulations in biosynthesis, cell adhesion, intracellular matrix binding such as fibrinogen-binding-related protein (Figure 5.6, Table 5.6).

The human associated strain Wright internalized at reduced oxygen, showed limited toxin production, and *agr* activity, and a few fermentative metabolic genes *fdaB, pflA,B* and thioredoxin (Figure 5.7, Table 5.7). This was coupled with increased expression of environmentally-sensitive alternative transcriptional regulators. Similarly, expression of hydroperoxide reductase (*ahpC*), another hydrogen peroxide scavenger, indicated that catalase activity was compensated. Expression of the PerR regulon genes, *ahpC* and *katA* was consistent with reduced oxygen and persistence properties (Cosgrove et al., 2007). In addition, similar expression patterns have been shown to be induced by osmotic pressure and NaCL (Armstrong-Buisseret et al., 1995), and acid shock (Bore et al., 2007). These environmental sensing profiles were further supported by the activation of *sigB* ponteially through *merR* (Brown et al., 2003), and the universal down regulation in *agr* and *sarU* related invasive phase products such as toxins (Figure 5.8 and Table 5.8). Similarly, cessation of fermentative pathway and limitations in the large-conductance mechanosensitive channel (*mscL*, locus SAB1205c) (Kouwen et al., 2009) and NADH dehydrogenase strongly suggested down regulation of *walKR* system (Dubrac et al., 2007; Dubrac and Msadek, 2004). More importantly, *ccpA* controlled pathways such as GntR and members of AraC/XylS and GalR/LacI families were not active (Seidl et al., 2009; Perez-Rueda and Collado-Vides, 2000). These observations suggested a major shift in gene expression and that reduced oxygen might have promoted persistence properties evidenced by metabolic inactivity (Shompole et al., 2003; Wesson et al., 1998; (Proctor et al., 2006).

In summary, in this comprehensive thesis projects we were able to establish recent approaches and link molecular typing and basic mechanisms underlying adaptation and

selection of *S. aureus* lineage in the bovine mammary gland. We have evaluated and used *clfA*-based typing system and significantly facilitated host- and organ-specific grouping of *S. aureus* strains from different host species, and further distinguished the clonal nature of mammary specific types. The major features of the system were the rapidity, sensitivity, specificity, and high reproducibility as well as the discriminatory power which matched that of major molecular methods. In addition, we were able to, for the first time, identify and characterize mastitis-associated *S. aureus* genotypes and their distribution, as well as subclonal populations in different geographic locations in Canada. Furthermore, we have investigated changes in gene expression of mammary epithelial cell internalized mastitis- and human-associated *S. aureus* strains that were precisely identified and characterized for this purpose. Thus, we have shown, for the first time, the role of reduced oxygen in altering *S. aureus* gene expression patterns in the mammary cells, and have revealed novel regulatory pathways and subsets of genes that were co-ordinately either up or down regulated in each strain during invasion processes under normal and reduced oxygen levels. In addition, a large number of genes with unknown and hypothetical fucntions were identified. Future annotation projects would potentially reveal a precise network of regulatory processes that are involved in the mammary gland microenvironment enabling significant insights into the basic mechanism of host- and tissue-specialization as well as in the selection of suitable therapeutic candidates.

REFERENCES

Ahmed S, Meghji S, Williams R, Henderson B, Brock J, Nair S. 2001. *Staphylococcus aureus* fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria. Infect Immun 69: 2872-2877.

Armstrong-Buisseret L, Cole MB, Stewart GS. 1995. A homologue to the *Escherichia coli* alkyl hydroperoxide reductase AhpC is induced by osmotic upshock in *Staphylococcus aureus.* Microbiology 141:1655-1661.

Atkins K, Burman J, Chamberlain E, Cooper J, Poutrel B, Bagby S, Jenkins A, Feil E, van den Elsen J. 2008. *S. aureus* IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. Mol Immunol 45: 1600-1611.

Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K.2002. Genome and virulence determinants of high virulence community-acquired MRSA. Lancet. 359: 1819–1827.

Bancroft EA. 2007. Antimicrobial Resistance: It's Not Just for Hospitals *JAMA*. 298:1803-1804.

Belcheva A, Golemi-Kotra D. 2008. A close-up view of the VraSR two-component system. A mediator of *Staphylococcus aureus* response to cell wall damage. J Biol Chem. 283:12354-64.

Bore E, Langsrud S, Langsrud O, Rode TM, Holck A. 2007. Acid-shock responses in *Staphylococcus aureus* investigated by global gene expression analysis. Microbiology 153:2289-2303.

Bronner S, Monteil H, Prévost G. 2004. Regulation of virulence determinants in *Staphylococcus aureus*: complexity and applications. FEMS Microbiol Rev 28:183-200.

Brown NL, Stoyanov JV, Kidd SP, Hobman JL. 2003. The MerR family of transcriptional regulators. FEMS Microbiol Rev. 27:145-63.

Cheng A, Kim H, Burts M, Krausz T, Schneewind O, Missiakas D. 2009. Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. FASEB J. 23:3393-404.

Cheung A, Nishina K, Trotonda M, Tamber S. 2008. The SarA protein family of *Staphylococcus aureus*. Int J Biochem Cell Biol 40: 355-361.

Cosgrove K, Coutts G, Jonsson IM, Tarkowski A, Kokai-Kun JF, Mond JJ, Foster SJ. 2007. Catalase (KatA) and alkyl hydroperoxide reductase (AhpC) have compensatory roles in peroxide stress resistance and are required for survival, persistence, and nasal colonization in *Staphylococcus aureus*. J Bacteriol. 189:1025-1035.

Criso´ stomo MI, Westh H, Tomasz A, Chung M, Oliveira DC, de Lencastre H,2001. The evolution of methicillin resistance in *Staphylococcus aureus*: similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. Proc. Natl. Acad. Sci. U.S.A.98, 9865–9870.

Ding Y, Onodera Y, Lee JC, Hooper DC. 2008. NorB, an efflux pump in *Staphylococcus aureus* strain MW2, contributes to bacterial fitness in abscesses. J Bacteriol 190 21:7123-9.

Dordet-Frisoni E, Dorchies G, De Araujo C, Talon R, Leroy S. 2007. Genomic diversity in *Staphylococcus xylosus*. Appl Environ Microbiol 73: 7199-7209.

Dubrac S, Msadek T. 2004. Identification of genes controlled by the essential YycGlYycF two-component system of *Staphylococcus aureus*. J Bacteriol 186:1175-1181.

Dubrac S, Boneca IG, Poupel O Msadek T. 2007. New insights into the WaIK/WalR (YycGlYycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. J Bacteriol 189:8257-8269.

Fitzgerald J, Meaney W, Hartigan P, Smyth C, Kapur V. 1997. Fine-structure molecular epidemiological analysis of *Staphylococcus aureus* recovered from cows. Epidemiol Infect 119: 261-269.

Foster TJ and Hook M. 1998. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 6: 484-488.

Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. Arac/XylS family of transcriptional regulators. Microbiol Mol Biol Rev. 61:393-410.

Highlander SK, Hultén KG, Qin X, Jiang H, Yerrapragada S, Mason EO Jr, Shang Y, Williams TM, Fortunov RM, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox GE, Cardenas AM, Muzny DM, Hemphill L, Ding Y, Dugan S, Blyth PR, Buhay CJ, Dinh HH, Hawes AC, Holder M, Kovar CL, Lee SL, Liu W, Nazareth LV, Wang Q, Zhou J, Kaplan SL, Weinstock GM.2007. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. BMC Microbiol 7: 99.

Higgins J, Loughman A, van Kessel K, van Strijp J, Foster T. 2006. Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. FEMS Microbiol Lett 258: 290-296.

Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabbinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J.2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci USA101:9786–97891.

Josefsson E, Kubica M, Mydel P, Potempa J, Tarkowski A. 2008. In vivo sortase A and clumping factor A mRNA expression during *Staphylococcus aureus* infection. Microb Pathog 44: 103-110.

Kapur V, Sischo W, Greer R, Whittam T, Musser J. 1995. Molecular population genetic analysis of *Staphylococcus aureus* recovered from cows. J Clin Microbiol 33: 376-380.

Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Zell ER, Fosheim GE, McDougal LK, Carey RB, Fridkin SK; Active Bacterial Core surveillance (ABCs) MRSA Investigators. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA* 298:1763-1771.

Kouwen TR, Trip EN, Denham EL, Sibbald MJ, Dubois JY, van Dijl JM. 2009. The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168. Antimicrob Agents Chemother 53:4702-4711

Lindsay JA, Moore CE, Day NP, Peacock SJ, Witney AA, Stabler RA, Husain SE, Butcher PD and Hinds J. 2006. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J Bacteriol 188: 669–676.

McGavin MJ, Zahradka C, Rice K Scott JE. 1997. Modification of the *Staphylococcus aureus* fibronectin binding phenotype by V8 protease. Infect Immun 65: 2621-2628.

Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG.1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA. 95:3140–3145.

Nickerson NN, Joag V, McGavin MJ. 2008. Rapid autocatalytic activation of the M4 metalloprotease aureolysin is controlled by a conserved N-terminal fungalysin-thermolysin-propeptide domain. Mol Microbiol 69: 1530–1543.

Nickerson NN, Prasad L, Jacob L, Delbaere LT, McGavin MJ. 2007. Activation of the SspA serine protease zymogen of *Staphylococcus aureus* proceeds through unique variations of a trypsinogen-like mechanism and is dependent on both autocatalytic and metalloprotease mtalloprotease-specific processing. J Biol Chem 282: 34129–34138.

Palmqvist N, Josefsson E, Tarkowski A. 2004a. Clumping factor A-mediated virulence during *Staphylococcus aureus* infection is retained despite fibrinogen depletion. Microbes Infect 6: 196-201.

Palmqvist N, Patti J, Tarkowski A, Josefsson E. 2004b. Expression of staphylococcal clumping factor A impedes macrophage phagocytosis. Microbes Infect 6: 188-195.

Perez-Rueda E, Collado-Vides J. 2000. The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. Nucleic Acids Research 28: 1838-1847.

Proctor RA, von Eiff C, Kahl BC, Becker K, McNamara P, Herrmann M, and Peters G. 2006. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. *Nature Rev. Microbiol*. 4: 295-305.

Rzychon M, Sabat A, Kosowska K, Potempa J, Dubin A. 2003. Staphostatins: an expanding new group of proteinase inhibitors with a unique specificity for the regulation of Staphopains, *Staphylococcus spp*. cysteine proteinases. Mol Microbiol 49**:** 1051–1066.

Sabria-Leal M, Morthland VH, Pedro-Botet ML, Sopena N, Gimenez-Perez M, Branchini ML, Pfaller MA.1994. Molecular epidemiology for local outbreaks of methicillin resistant *Staphylococcus aureus* (MRSA). The need for several methods. Eur. J. Epidemiol. 10, 325–330.

Saei H, Ahmadi M, Mardani K, Batavani R. 2009. Molecular typing of *Staphylococcus aureus* isolated from bovine mastitis based on polymorphism of the coagulase gene in the north west of Iran. Vet Microbiol 137: 202-206.

Seidl K, Müller S, François P, Kriebitzsch C, Schrenzel J, Engelmann S, Bischoff M, Berger-Bächi B. 2009. Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus.* BMC Microbiol. 18: 9:95.

Sendi P, Proctor RA. 2009. *Staphylococcus aureus* as an intracellular pathogen: the role of small colony variants. Trends Microbiol 17:54-88.

Shompole S, Kim TH, Linda E L, Katarzyna D, Gregory AB and Kenneth WB. 2003. Biphasic intracellular expression of *Staphylococcus aureus* virulence factors and evidence for Agr-mediated diffusion sensing. Mol. Microbiol. 49:919–927.

Skaar EP, Schneewind O. 2004. Iron-regulated surface determinants (Isd) of *Staphylococcus aureus*: stealing iron from heme. Microbes and Infection 6:390-397.

Ster C, Gilbert F, Cochard T, Poutrel B. 2005. Transcriptional profiles of regulatory and virulence factors of *Staphylococcus aureus* of bovine origin: oxygen impact and strain-to-strain variations. Mol Cell Probes 19: 227-235.

Su C, Herbelin C, Frieze N, Skardova O, Sordillo L. 1999. Coagulase gene polymorphism of *Staphylococcus aureus* isolates from dairy cattle in different geographical areas. Epidemiol Infect 122: 329-336.

Takeuchi F, Watanabe S, Baba T, Yuzawa H, Ito T, Morimoto Y, Kuroda M, Cui L, Takahashi M, Ankai A, Baba S, Fukui S, Lee JC, Hiramatsu K. 2005. Whole-genome sequencing of *Staphylococcus haemolyticus* uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. J Bacteriol 187: 7292-7308.

Taylor BL, Zhulin IB. 1999. PAS domains: internal sensors of oxygen, redox potential and light. Microbiol Mol Bioi Rev 63:479-506.

Tuchscherr L, Buzzola F, Alvarez L, Lee J, Sordelli D. 2008. Antibodies to capsular polysaccharide and clumping factor A prevent mastitis and the emergence of unencapsulated and small-colony variants of *Staphylococcus aureus* in mice. Infect Immun 76: 5738-5744.

Vaudaux P, Francois P, Bisognano C, Kelley WL, Lew DP, Schrenzel J, Proctor RA, McNamara PJ, Peters G, Von Eiff Ch. 2002. Increased expression of clumping factor and fibronectin-binding proteins by *hemB* mutants of *Staphylococcus aureus* expressing small colony variant phenotypes. Infect. Immun 70:5428–5437.

Visai L, Yanagisawa N, Josefsson E, Tarkowsk A, Pezzali I, Rooijakkers SHM. Foster TJ, Speziale P. 2009. Immune evasion by *Staphylococcus aureus* conferred by iron-regulated surface determinant protein IsdH. Microbiology 155:667-679.

Villari P, Iacuzio L, Torre I, Scarcella A, 1998. Molecular epidemiology as aneffective tool in the surveillance of infections in the neonatal intensive care unit. J. Infect. 37, 274–281.

Weinrick B, Dunman PM, McAleese F, Murphy E, Projan SJ, Fang Y, Novick RP.2004. Effect of mild acid on gene expression in *Staphylococcus aureus*. J Bacteriol. 186:8407-23.

Wesson C, Liou L, Todd K, Bohach G, Trumble W, Bayles K. 1998. *Staphylococcus aureus* Agr and Sar global regulators influence internalization and induction of apoptosis. Infect Immun 66: 5238-5243.

Wolz C, McDevitt D, Foster T, Cheung A. 1996. Influence of agr on fibrinogen binding in *Staphylococcus aureus* Newman. Infect Immun 64: 3142-3147.

Wolz, C, Goerke C, Landmann R, Zimmerli W, Fluckiger U. 2002. Transcription of clumping factor A in attached and unattached *Staphylococcus aureus* in vitro and during device-related infection. Infect. Immun 70:2758–2762.

Yarwood JM, Paquette KM, Tikh IB, Volper EM, Greenberg EP. 2007. Generation of virulence factor variants in *Staphylococcus aureus* biofilms. J Bacteriol. 189:7961-7967.

Yesilkaya H, Spissu F, Carvalho SM, Terra VS, Homer KA, Benisty R, Porat N, Neves AR, Andrew PW. 2009. Pyruvate formate lyase is required for pneumococcal fermentative

metabolism and virulence. Infect Immun. [Epub ahead of print doi:10.1128/IAI.00178-09].