
Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/13835/1/OTOKUNEFOR_KOME_PHD_THESIS_DEC_2013.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see: http://eprints.nottingham.ac.uk/end_user_agreement.pdf

For more information, please contact eprints@nottingham.ac.uk
Development and validation of novel methods for the study of *Staphylococcus aureus* PVL strains

Kome Otokunefor, BSc, MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

©December 2013
DECLARATION

I declare that the work presented in this thesis is my own unless otherwise stated.

Signed.................................................................
ABSTRACT

Since the initial association of the Panton Valentine leucocidin (PVL) toxin with highly virulent strains of *Staphylococcus aureus* found in the community, a firm epidemiological link has been established between the PVL encoding genes and community-acquired strains of both meticillin resistant (CA-MRSA) and susceptible (CA-MSSA) *Staphylococcus aureus*. While most research has predominantly concentrated on the genotyping of CA-MRSA strains, PVL-MSSA appear to pose an increasing public health risk. Currently though, there exists a dearth of epidemiological data on PVL-MSSA strains, particularly with regard to the *lukS* and *lukF* genes which encode the PVL toxin.

This first aim of the present study therefore was to explore the genetic diversity of a group of PVL-MSSA clinical local isolates in order to contribute to the limited current data and provide insight into the evolution and emergence of PVL-MRSA isolates. In addition, as current typing systems are cumbersome, time consuming and expensive, this present study was also aimed at the development of a rapid high resolution melt (HRM) typing system for the characterisation of PVL-positive isolates. The PVL toxin is encoded for by two highly conserved adjacent genes (*lukF* and *lukS*) which are co-transcribed. Variations in these genes correlate with a strain’s genotype. Therefore, the present study set out to genotype isolates based on the four major SNPs at positions 527 and 663 of the *lukS* gene and 1396 and 1729 of the *lukF* gene. The final aim of the present study was the development of an enzyme linked immunosorbent assay (ELISA) system (for the detection and quantification of both PVL and alpha haemolysin) that has potential application in clinical diagnosis and as a research tool.

Characterisation of a collection of UK PVL-MSSA isolates by MLST and spa typing which is presented in the present study, showed a genetic similarity to circulating PVL-MRSA strains, with 94.7% of the isolates belonging to CA-MRSA related genetic backgrounds (ST1, ST22, ST30, ST772 and ST88). Three novel spa types (t6642, t6643 and t6769) and a novel ST (ST1518) were however detected in this population. Furthermore, the presence of identical PVL phages and haplotypes in the PVL-MSSA isolates to those previously described in PVL-MRSA isolates point
strongly at the role these strains may play as precursors of emerging lineages of clinical significance.

The HRM assay developed in this study was able to accurately genotype PVL-positive isolates based on the double allelic variations in both the $lukS$ and $lukF$ genes. The high degree of sensitivity of this technique was clearly demonstrated by its ability to differentiate between the $lukS\ A_{527}/G_{663}$ and $G_{527}/T_{663}$ genotypes which theoretically should have the same melt temperature. Despite the issues in data interpretation, which arose following attempts to improve the discrimination of this technique by the addition of a third locus ($spa$), the technique still showed potential as a useful tool for the rapid genotyping of PVL-positive isolates.

While HRM was useful in rapidly detecting and genotyping PVL-positive isolates, the actual level of protein production of both PVL and HLA toxins could only be determined following the development and validation in the present study of a simple competitive ELISA platform which exploited the high affinity biotin/streptavidin interaction to improve sensitivity. This technique would be especially useful in settings which lack the specialised equipment required for genetic studies like HRM.

In summary, in addition to contributing to the limited epidemiological information about PVL-MSSA strains and demonstrating a clear role for these strains in the evolution of PVL-MRSA strains, the present study has developed two distinct methods to aid the study of $S.\ aureus$ PVL producing strains which are becoming a significant healthcare problem worldwide. The present study will contribute to our understanding of these strains and to the development of intervention strategies to curb their spread and threat.
ACKNOWLEDGEMENTS

This piece of work and indeed the last three years would not have been possible without the input, guidance, help and support of so many groups of people.

So here is a big thank you to all those who in one way or the other made this possible!

My supervisors Prof. Richard James and Dr. Alan Cockayne
For unparalleled supervision and invaluable support

The people of 2B Gambia-Street

Staph Group
Naija folks
Friends
Loughborough haven
Commonwealth Scholarship commission
Luton clan
Grace Family
CBS colleagues

boRgel
For friendship and great times

The King who holds my world in His hand!
TABLE OF CONTENTS

Declaration ...........................................................................................................1
Abstract ................................................................................................................2
Acknowledgements ............................................................................................4
Table of Contents ...............................................................................................5
List of Figures ....................................................................................................10
List of Tables .....................................................................................................14
List of Appendices .............................................................................................16
Abbreviations ....................................................................................................17

1 Introduction and Literature Review .........................................................21

1.1 Introduction .................................................................................................21
1.2 Staphylococcus aureus ...............................................................................23
1.3 Virulence Determinants .............................................................................26
1.4 Antibiotic Resistance ..................................................................................30
1.4.1 Meticillin And Development Of Meticillin Resistance .......................31
1.4.2 Mechanism of Meticillin Resistance .....................................................35
1.4.2.1 meca Mediated Meticillin Resistance .............................................35
1.4.2.2 Non-meca Mediated Resistance .....................................................39
1.5 Community-acquired MRSA .....................................................................40
1.5.1 Characteristics of CA-MRSA strains ...................................................42
1.6 Panton-Valentine leucocidin ......................................................................44
1.6.1 Link to CA-MRSA ................................................................................44
1.6.2 PVL Origins ............................................................................................49
1.6.3 Genetics of PVL ....................................................................................51
1.6.4 PVL phages ............................................................................................54
1.6.5 PVL Controversy ...................................................................................54
1.6.6 PVL Expression Studies .........................................................................58
2.12.2 Protein Separation and Visualisation

2.12.1 Preparation of E. coli cell lysates by sonication

2.12.2 Cell Fractionation

2.11 Methods used in DNA manipulation

2.11.1 A-tailing

2.11.2 DNA Ligation

2.11.3 Restriction Enzyme (RE) Digestion

2.11.4 Preparation of Competent E. coli

2.11.5 Transformation of E. coli by electroporation

2.10 DNA Electrophoresis

2.10.1 Purification of PCR products

2.10.2 DNA Sequencing

2.9 Polymerase Chain Reaction (PCR)

2.8 Primer Design and Synthesis

2.7 DNA quantification

2.6 Plasmid DNA Extraction from E. coli

2.5 Genomic DNA Extraction

2.4 Growth of Staphylococcus aureus and Escherichia coli

2.3 PVL-positive test isolates

2.2 Chemicals

2.1 Media

1.11 Aims and Objectives

1.10 Conclusion

1.9 Restriction-Modification Systems

1.8 Typing approaches used to study evolution of S. aureus

1.8.1 eBURST

1.8.2 MRSA Evolution

1.7 S. aureus typing

1.7.4 High Resolution Melt Analysis

1.7.3 spa Typing

1.7.2 Multilocus Sequence Typing (MLST)

1.7.1 PFGE

1.7 DNA Electrophoresis

1.6.8 Phenol Soluble Modulins (PSMs)

1.6.7 Alpha Haemolysin/Toxin

1.6 Cell Fractionation

1.5 Preparation of Transformation of

1.4 Preparation of Competent

1.3 Restriction Enzyme (RE) Digestion

1.2 Preparation of Competent E. coli

1.1 Transformation of E. coli by electroporation

1. Protein Separation and Visualisation

1.0 Aims and Objectives

2 General Materials and Methods

83
3 Molecular characterization and PVL typing of community-acquired meticillin-sensitive \textit{S. aureus} clinical isolates .......................... 95

3.1 Introduction .................................................................................................................. 95
3.2 Materials and Methods .................................................................................................. 99
3.2.1 Bacterial Isolates ..................................................................................................... 99
3.2.2 Strain Characterisation ........................................................................................... 99
3.2.2.1 \textit{spa} typing ........................................................................................................ 100
3.2.2.2 Multilocus Sequence Typing (MLST) ............................................................... 103
3.2.2.3 PVL characterization and PVL phage typing ..................................................... 103
3.3 Results .......................................................................................................................... 109
3.3.1 Characteristics of PVL-MSSA ................................................................................. 109
3.3.2 \textit{spa} typing of test isolates ................................................................................... 111
3.3.3 Multilocus Sequence Typing .................................................................................. 120
3.3.4 SNP analysis of the PVL-encoding genes ............................................................. 122
3.3.5 PVL Phage typing .................................................................................................... 124
3.3.6 Summary of Results ............................................................................................... 131
3.4 Discussion ..................................................................................................................... 133

4 Development and validation of a high resolution melt based method for the rapid typing/identification of PVL-positive strains of \textit{S. aureus} .................................................................................................................. 143

4.1 Introduction .................................................................................................................... 143
4.1.1 High Resolution Melt (HRM) ................................................................................ 145
4.2 Materials and Methods .................................................................................................. 149
4.2.1 Primers..................................................................................................................... 149
4.2.2 Standard \textit{S. aureus} Isolates ................................................................................ 149
4.2.3 HRM ....................................................................................................................... 152
4.3 Results .......................................................................................................................... 154
4.3.1 \textit{lukSF-PV} locus HRM Analysis ......................................................................... 154
4.3.1.1 Pilot study to characterise a subset of isolates using HRM analysis .......... 158
4.3.2 \textit{spa} locus HRM analysis ....................................................................................... 160
4.3.2.1 Test strain characteristics ................................................................................. 160
4.3.2.2 HRM analysis of isolates of differing \textit{spa} type .......................................... 161
4.3.2.3 \textit{spa} HRM Data Interpretation ........................................................................ 163
4.3.2.4 Effect of well position on results of HRM analysis ...................................... 168
5 Development and Validation of an Enzyme-Linked ImmunoSorbent Assay for the detection and quantification of PVL and HLA toxins ........................................................................................................................................222

5.1 Introduction ..................................................................................................................222
5.1.1 Enzyme-Linked Immunosorbent Assay (ELISA) ................................................224
5.2 Materials and Methods ..............................................................................................230
5.2.1 Pre-protein expression and purification steps .....................................................230
5.2.1.1 Background ....................................................................................................230
5.2.1.2 Construction of pET-21d(+)-hlaH35L plasmid expression construct ........231
5.2.1.3 pET-21d(+)-lukF expression construct .......................................................232
5.2.2 Protein expression .................................................................................................234
5.2.3 Protein Purification ...............................................................................................235
5.2.4 Protein Quantification .........................................................................................235
5.2.5 Protein Biotinylation ..........................................................................................236
5.2.5.1 ELISA to evaluate protein biotinylation .......................................................236
5.2.6 Competitive ELISA ............................................................................................237
5.3 Results .......................................................................................................................239
5.3.1 Confirmation of hla-H35L sequence prior to protein expression .................239
5.3.2 Pilot studies to confirm HLA-H35L and LukF protein expression .................242
5.3.3 Affinity Protein Purification .................................................................................245
5.3.4 Modification of HLA-H35L purification procedure .......................................247
5.3.5 Use of purified recombinant LukF and HLA-H35L proteins in competitive ELISA development .................................................................251
5.3.5.1 Protein Biotinylation .....................................................................................251
5.3.6 Competitive ELISA exploring competitive ability of biotinylated proteins .. 254
Pilot studies using spiked culture media in competitive ELISA ............ 256
Validation of ELISA ........................................................................... 263
Generation of Standard Inhibition Curves for protein quantification using the competitive ELISAs ......................................................... 263
Application of Competitive ELISA to a panel of S. aureus isolates .... 267

Discussion ............................................................................................. 271
ELISA protocol .................................................................................... 271
Recombinant protein production .......................................................... 275
Summary ............................................................................................... 278

General Discussion .............................................................................. 280
Overview .............................................................................................. 280
Novel typing techniques ....................................................................... 289
Future Work .......................................................................................... 292
Conclusions ........................................................................................ 295
References ............................................................................................ 296
Appendices ............................................................................................ 348
LIST OF FIGURES

Figure 1.1: Illustration of the chemical formula of penicillin (A) and meticillin (B) ... 32

Figure 1.2: Schematic representation of variation in the eight SCCmec types described by 2009 showing the different composition and arrangement of mec (in pink) and ccr gene complexes (in blue) ................................................................. 38

Figure 1.3: Schematic representation of the lukSF-PV locus ........................................ 51

Figure 1.4: Illustration of the spa gene showing the polymorphic X region (Xr) used in spa typing and the positioning of the primer sequences used for both amplification and sequencing ................................................................. 69

Figure 3.1: Hypothetical pathways for the evolution of PVL-MRSA .................. 97

Figure 3.2: A representative example of expected product sizes for PVL phage amplification reactions ................................................................. 107

Figure 3.3: Example of a noisy sequence chromatogram generated following sequencing of the spa amplicons purified with a QiaQuick PCR purification kit .... 112

Figure 3.4: Representative gel images indicating purity of the spa gene fragments following amplicon purification by gel extraction ................................................................. 113

Figure 3.5: Example of a clean sequence chromatogram generated following sequencing of the spa amplicons purified with a QiaQuick gel extraction kit ........ 114

Figure 3.6: Representation of sequence analysis and the spa type allocation process for TS13 ................................................................. 116

Figure 3.7: Depiction of the relationship between the new spa types described in the present study and their existing clonal prototypes as defined by the BURP algorithm ................................................................. 119

Figure 3.8: Representative gel images depicting the effect of DMSO on generation of amplicons for PVL phage typing ................................................................. 125

Figure 3.9: Agarose gel image showing examples of non-specific amplification products following PVL phage typing ................................................................. 127

Figure 3.10: Agarose gel image depicting lack of expected amplification products following PVL-phage typing of an ST772 isolate ................................................................. 130
Figure 4.1: Process flowchart for sequence based typing systems ........................................... 144

Figure 4.2: Representation of different melt curves generated during the HRM analysis of four S. aureus NARSA isolates. ........................................................................................................ 155

Figure 4.3: Automatic grouping of NRS158 and NRS162 as S2 and S1 variants (circled) respectively................................................................. 157

Figure 4.4: Normalised and temperature shifted difference curves of five S. aureus isolates representing three different spa types using temperature shifts of 5°C or 0°C. ........................................................................................................ 162

Figure 4.5: Normalised and zero temperature shifted difference curves of five S. aureus test isolates representing three different spa types ........................................................................................................ 164

Figure 4.6: Automatic classification of normalised and zero temperature shifted difference curves of five S. aureus test isolates into three groups ........................................................................................................ 165

Figure 4.7: High resolution melt profile of five S. aureus isolates with identical t005 spa types ........................................................................................................ 167

Figure 4.8: Aliquot positions used to assay for effects of well position on HRM profiles ........................................................................................................ 168

Figure 4.9: Variable amplification profiles generated for aliquots of single master mixes for TS6 and TS9 analysed at different well positions of the 96-well plate. .... 170

Figure 4.10: Normalised and zero temperature shifted difference curves of a single sample (TS6) analysed at different positions of the 96-well plate................................. 171

Figure 4.11: Normalised and automatic temperature shifted difference curve clustering for aliquots of single master mixes for TS6 and TS9 analysed at different well positions of the 96-well plate........................................................................................................ 172

Figure 4.12: Normalised and zero temperature shifted difference curves of S. aureus test isolates based on the proposed criteria for HRM data analysis................................. 175

Figure 4.13: Representative Normalised and automatic temperature shifted difference curve profiles of the lukS locus ........................................................................................................ 179

Figure 4.14: Melt curves of lukS locus showing lower melt peaks of negative and blank samples ........................................................................................................ 180

Figure 4.15: Amplification curves showing variable crossing points generated from HRM analysis of lukS in a “blinded” panel of isolates ........................................................................................................ 182
Figure 4.16: Effect of automatic (A) and adjusted premelt (B) values on the consistency of HRM profiles for duplicate assay wells for isolates 4 and 9. .......... 184

Figure 4.17: HRM spa profiles for 20 isolates tested in triplicate......................... 187

Figure 4.18: HRM spa profiles generated for group A (S1F1) isolates....................... 188

Figure 4.19: HRM spa profiles generated for group C (S2F1) isolates....................... 189

Figure 4.20: HRM profiles of group C (S2F1) subgroup 1 isolates produced under varying conditions designed to clarify the genotype of closely related isolates........... 191

Figure 4.21: HRM profiles of group C (S2F1) subgroup 2 isolates produced under varying conditions designed to clarify the genotype of closely related isolates......... 193

Figure 4.22: HRM spa profiles generated for group D (S2F2) isolates......................... 194

Figure 4.23: HRM spa profiles of group D (S2F2) subgroup 3 only. ....................... 195

Figure 4.24: Subgroup D3 (S2F2) isolates showing complex HRM difference curves ................................................................................................................ 196

Figure 4.25: Normalised and automatic temperature shifted difference curves of lukS locus with identical DNA composition but varying SNP location ...................... 211

Figure 4.26: Illustration of across plate temperature variation observed in the different HRM platforms ............................................................................................................ 219

Figure 5.1: Illustration of the direct ELISA process.................................................. 226

Figure 5.2: Representations of Indirect (A) and Sandwich (B) ELISA processes.... 228

Figure 5.3: Map of the pET-21d(+) plasmid expression vector................................. 233

Figure 5.4: Flowchart depicting the two major ELISA protocols used in the present study. .................................................................................................................. 238

Figure 5.5: DNA sequence alignment of hla-H35L in pKO1 and the S. aureus Newman hla ..................................................................................................................... 240

Figure 5.6: Protein sequence alignment of the HLA_NWMN and............................... 241

Figure 5.7: SDS-PAGE analysis of recombinant HLA-H35L protein expression at 37°C in E. coli BL21/DE3pKO2. ...................................................................................... 243

Figure 5.8: SDS-PAGE analysis of recombinant LukF protein expression at 30°C or 37°C in E. coli BL21/DE3pKO3. ...................................................................................... 244
Figure 5.9: SDS-PAGE analysis of recombinant LukF (top panel) and HLA-H35L (bottom panel) expressed in *E. coli* BL21/DE3 and purified using the Novagen His Bind Resin.

Figure 5.10: SDS-PAGE analysis of recombinant HLA-H35L expressed in *E. coli* BL21/DE3 and purified using the BioRad® Biologic LP affinity chromatographic system.

Figure 5.11: SDS-PAGE analysis of purified and dialysed recombinant HLA-H35L.

Figure 5.12: Determination of the biotinylation efficiency of the recombinant HLA-H35L and LukF proteins obtained in two separate biotinylation reactions.

Figure 5.13: Pilot competitive ELISA results comparing the ability of two batches of biotinylated HLA-H35L and LukF to inhibit binding of unlabelled proteins.

Figure 5.14: Results of competitive ELISA using various culture media (BHI, LB and CYGP) spiked with HLA-H35L or LukF.

Figure 5.15: Effects of β-glycerophosphate and CYGP dilution on competitive ELISA results.

Figure 5.16: Specificity testing of LukF ELISA using CYGP culture supernatants of PVL-negative *S. aureus* isolates.

Figure 5.17: Standard inhibition calibration curves for HLA-H35L and LukF proteins generated using GraphPad Prism.

Figure 6.1: Incidence of reported MRSA bacteraemia in England (1993 – 2008).

Figure 6.2: Incidence of reported *S. aureus* bacteraemia in the UK (2002 – 2011).
LIST OF TABLES

Table 1.1: Spectrum of diseases associated with S. aureus ............................................ 25
Table 1.2: Composition of the eight SCCmec types described by 2009 ......................... 37
Table 1.3: Rising incidence of PVL-positive S. aureus isolates in the UK ...................... 47
Table 1.4: Relationship between lukSF-PV sequence variants and evolutionary origins ........................................................................................................................................................................... 53
Table 2.1: NUH NHS Trust PVL-positive S. aureus clinical isolates ............................. 85
Table 2.2: PVL-positive S. aureus isolates from the University of Nottingham S. aureus strain collection .................................................................................................................................................................................. 86
Table 3.1: Primers Used In DNA amplification in the present study .............................. 102
Table 3.2: Primers used for detecting and typing PVL phages ...................................... 106
Table 3.3: Key to PVL phage typing .................................................................................. 108
Table 3.4: Clinical characteristics of Panton-Valentine leucocidin positive S. aureus test isolates .......................................................................................................................................................................................... 110
Table 3.5: Results of spa typing ......................................................................................... 118
Table 3.6: MLST allelic profiles and corresponding STs .................................................. 121
Table 3.7: Single nucleotide polymorphisms detected in the lukSF-PV locus of the study population .......................................................................................................................................................................................... 123
Table 3.8: Results of PVL-phage typing of the clinical isolates ...................................... 129
Table 3.9: Molecular characterization of 19 PVL-positive S. aureus clinical isolatesa .................................................................................................................................................................................................................................. 132
Table 4.1: The four major SNP combinations used to define PVL isoforms ............ 147
Table 4.2 Allelic variation of standard isolates used for the development of the lukSF-PV HRM assay .................................................................................................................................................................................................................................................. 150
Table 4.3: spa type and repeat sequence of isolates used for the development of spa HRM ........................................................................................................................................................................................................................................ 151
Table 4.4: HRM profiles of isolates generated following HRM based typing ............ 159
Table 4.5: List of isolates used to test the optimised spa HRM ............................... 175

Table 4.6: Relevant characteristics of isolates used to validate the HRM assay by combining PVL and spa type data. ................................................................. 177

Table 4.7: Summary of lukS and lukF HRM results of 40 blinded isolates............. 185

Table 4.8: Compilation of lukSF-PV and spa HRM results to create unique code for each genotype................................................................. 198

Table 4.9: Comparison of experimentally generated HRM genotypes with predicted results .................................................................................. 199

Table 4.10: Variations from expected PVL genotypes detected via sequencing...... 203

Table 4.11: Linking of blinded isolate identity to experimentally determined isolate identities .................................................................................. 204

Table 4.12: spa HRM predictive ability ................................................................. 206

Table 4.13: Accuracy of spa HRM groupings compared to known spa types ....... 208

Table 4.14: Single nucleotide polymorphisms in the 4 different test lukS sequences in comparison to a parent sequence* ................................................................. 210

Table 4.15: Variable characteristics of the lukS test sequences ....................... 212

Table 5.1: Correlation between actual and interpolated protein concentrations...... 266

Table 5.2: Detection and Quantification of LukF and HLA proteins using the newly developed ELISA platforms ................................................................. 269

Table 5.3: Comparison of LukF and HLA protein production for selected isolates.. 270
LIST OF APPENDICES

Appendix 1  Media Composition..........................................................348
Appendix 2  Buffer/Chemical Compositions........................................350
Appendix 3  Identity of Blinded Isolates............................................353
Appendix 4  Demonstration of lack of haemolytic activity of HLA-H35L on blood agar plate.......................................................354
Appendix 5  Raw ELISA results..........................................................355
Appendix 6  JCM Publication.............................................................356
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>µl</td>
<td>Microliter</td>
</tr>
<tr>
<td>A$_{260}$</td>
<td>Absorbance at 260nm</td>
</tr>
<tr>
<td>Ab</td>
<td>Antibody</td>
</tr>
<tr>
<td>ABTS</td>
<td>2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)</td>
</tr>
<tr>
<td>Ag</td>
<td>Antigen</td>
</tr>
<tr>
<td>agr</td>
<td>Accessory gene regulator</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BA</td>
<td>Biotin-avidin</td>
</tr>
<tr>
<td>BHI</td>
<td>Brain Heart Infusion</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>BORSA</td>
<td>Borderline resistant <em>S. aureus</em></td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>BSAU</td>
<td>Biopolymer Synthesis and Analysis Unit</td>
</tr>
<tr>
<td>CA-MRSA</td>
<td>Community-acquired meticillin resistant <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>CA-MSSA</td>
<td>Community-acquired meticillin susceptible <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>CC</td>
<td>Clonal Complex</td>
</tr>
<tr>
<td>CCY</td>
<td>Casein hydrolysate-yeast extract containing</td>
</tr>
<tr>
<td>cfu</td>
<td>Colony forming unit</td>
</tr>
<tr>
<td>CHIPS</td>
<td>Chemotaxis inhibitory protein of <em>S. aureus</em></td>
</tr>
<tr>
<td>CoNS</td>
<td>Coagulase negative Staphylococci</td>
</tr>
<tr>
<td>Cp</td>
<td>Crossing point</td>
</tr>
<tr>
<td>CRISPRs</td>
<td>Clustered, regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytosine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>DLV</td>
<td>Double locus variant</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphates</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double stranded DNA</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxythymidine triphosphate</td>
</tr>
<tr>
<td>eBURST</td>
<td>Enhanced based upon related sequence type</td>
</tr>
<tr>
<td>Efb</td>
<td>Extracellular fibrinogen-binding protein</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-Linked Immunosorbent Assay</td>
</tr>
<tr>
<td>EMRSA</td>
<td>Epidemic MRSA</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>GEI</td>
<td>Genomic Island</td>
</tr>
<tr>
<td>GP</td>
<td>General Practice</td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>Water</td>
</tr>
<tr>
<td>HA-MRSA</td>
<td>Hospital-acquired meticillin resistant <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>HLA</td>
<td>Alpha toxin</td>
</tr>
<tr>
<td>HPA</td>
<td>Health Protection Agency</td>
</tr>
<tr>
<td>HRM</td>
<td>High resolution melt</td>
</tr>
<tr>
<td>HRP</td>
<td>Horse Radish Peroxidase</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilised metal affinity chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>KC</td>
<td>Keratinocyte-derived chemokine</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>LA-MRSA</td>
<td>Livestock associated meticillin resistant <em>S. aureus</em></td>
</tr>
<tr>
<td>LAB</td>
<td>Labelled avidin-biotin</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>LP</td>
<td>Low Pressure</td>
</tr>
<tr>
<td>LOD</td>
<td>Limit of detection</td>
</tr>
<tr>
<td>LOQ</td>
<td>Limit of quantification</td>
</tr>
<tr>
<td>mAB</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>MGE</td>
<td>Mobile genetic elements</td>
</tr>
<tr>
<td>mgr</td>
<td>Multiple gene regulator</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimum inhibitory concentration</td>
</tr>
<tr>
<td>MIP-2</td>
<td>Macrophage inflammatory protein-2</td>
</tr>
<tr>
<td>ml</td>
<td>Millilitre</td>
</tr>
<tr>
<td>MLEE</td>
<td>Multilocus enzyme electrophoresis</td>
</tr>
<tr>
<td>MLST</td>
<td>Multilocus sequence typing</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MRSA</td>
<td>Meticillin resistant <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>MSSA</td>
<td>Meticillin susceptible <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NARSA</td>
<td>Network on Antimicrobial Resistance in <em>S. aureus</em></td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NGS</td>
<td>Next Generation Sequencing</td>
</tr>
<tr>
<td>NUH</td>
<td>Nottingham University Hospitals</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PAI</td>
<td>Pathogenicity island</td>
</tr>
<tr>
<td>PBP</td>
<td>Penicillin binding protein</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PBST</td>
<td>PBS-Tween</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PFGE</td>
<td>Pulsed field gel electrophoresis</td>
</tr>
<tr>
<td>PMN</td>
<td>Polymorphonuclear neutrophil</td>
</tr>
<tr>
<td>PSM</td>
<td>Phenol soluble modulin</td>
</tr>
<tr>
<td>PVL</td>
<td>Panton-Valentine leucocidin</td>
</tr>
<tr>
<td>RAPD</td>
<td>Random Amplification of Polymorphic DNA</td>
</tr>
<tr>
<td>RE</td>
<td>Restriction enzyme</td>
</tr>
<tr>
<td>RGB</td>
<td>Resolving gel buffer</td>
</tr>
<tr>
<td>RM</td>
<td>Restriction Modification</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>rot</td>
<td>Repressor of toxin</td>
</tr>
</tbody>
</table>
rpm  Revolution per minute
RS   Richard Spence
RSD  Relative signal difference
RT-PCR  Real time PCR
sae  S. aureus exoprotein expression
SaPI  Staphylococcal pathogenicity island
sar  Staphylococcal accessory gene regulator
SCC  Staphylococcal cassette chromosome
SCIN  Staphylococcal complement inhibitor
SDS  Sodium dodecyl sulphate
SGB  Stacking gel buffer
SLV  Single locus variant
SNP  Single nucleotide polymorphism
SPA  Staphylococcal protein A
SRU  Staphylococcus Reference Unit
SSL-7  Staphylococcal superantigen-like protein-7
ST  Sequence type
Taq  Thermus aquaticus
TEMED  Tetramethylethylenediamine
T_m  Melting Temperature
TS  Tim Sloan
TSB  Trypticase Soy Broth
TSST  Toxic shock syndrome toxin
U  Units
UV  Ultra violet
v/v  Volume per volume
WGS  Whole genome sequencing
w/v  Weight per volume
xpr  Extracellular protein regulator
Chapter One

1 Introduction and Literature Review

1.1 Introduction

In 2002, initial reports were made linking the cytotoxic Panton-Valentine leucocidin (PVL) toxin specifically with community-acquired *S. aureus* (CA-MRSA) as opposed to the hospital-acquired *S. aureus* (HA-MRSA) strains (Gillet et al., 2002). Subsequently, over the years, an epidemiological link has been firmly established between PVL encoding genes and community-acquired strains of both meticillin resistant (CA-MRSA) and susceptible (CA-MSSA) *S. aureus* strains (Chheng et al., 2009, Diep et al., 2004, Enany et al., 2007, Liassine et al., 2004, Linde et al., 2005, Miklasevics et al., 2004, Tsai et al., 2008, Vandenesch et al., 2003, Witte et al., 2005). These CA-MRSA isolates often cause severe infection in young healthy immunocompetent individuals and differ from HA-MRSA in antimicrobial susceptibility pattern, clonal origins and type of *mec* element. Furthermore, PVL encoding genes were found to be highly represented within this population. Up to 100% prevalence was observed among CA-MRSA strains as opposed to the 2% frequency noted in the general *S. aureus* population (Kilic et al., 2006, Kilic et al., 2008, Makgotlho et al., 2009, Prevost et al., 1995a). PVL toxin is encoded by two adjacent, co-transcribed genes, *lukF* and *lukS*. These genes which make up the highly conserved \(\approx 1.9\) kb *lukSF-PV* locus are carried by several temperate phages. (Kaneko et al., 1998, Kaneko et al., 1997, Narita et al., 2001, Baba et al., 2002, Diep et al., 2006b, Ma et al., 2006, Ma et al., 2008, Takano et al., 2008b, Berglund et al., 2008b, Otter et al., 2010). Despite the undeniable epidemiological link between PVL
encoding genes and CA-MRSA, the contribution of the PVL toxin in both the virulence and pathogenesis of this group of isolates has not yet been conclusively demonstrated.

Current research in this field has hence been geared towards answering this question. Most virulence studies involved the use of various animal and infection models but failed to provide actual evidence of toxin production (Labandeira-Rey et al., 2007, Voyich et al., 2006, Bubeck Wardenburg et al., 2008). A few studies which focused on analysing actual toxin production in infected tissues and in vitro, noted wide strain to strain variations in the amount of PVL toxin produced by clinical isolates (Badiou et al., 2008, Croze et al., 2009, Hamilton et al., 2007, Loughman et al., 2009, Said-Salim et al., 2005). In addition to the wide array of factors already noted to affect PVL expression levels (Bronner et al., 2000, Dumitrescu et al., 2008a), more recent findings (Wirtz et al., 2009) on differential transcription of PVL encoding genes from the øSa2mw phage, following mitomycin treatment in several S. aureus strains, appear to indicate an effect of staphylococcal genetic background on PVL expression. Over the years, genetic background has played a pivotal role in the evolution, spread and epidemiology of S. aureus particularly with respect to the acquisition of the meca element (Enright et al., 2002) and virulence genes (Diep et al., 2006a), the global spread of MRSA clones (Vandenesch et al., 2003), and agr type (Jarraud et al., 2002). The relationship between the genetic background of these strains and the levels of both PVL and alpha haemolysin toxins produced has however not yet been investigated. Strain typing to generate the necessary information on genetic background would be necessary to answer these questions. Despite the acceptable
discriminatory abilities of current typing methods, these methods can however be cumbersome, expensive and time consuming.

1.2 Staphylococcus aureus

Staphylococci

In 1881, Alexander Ogston a Scottish physician was the first person to describe a distinct form of ‘grouped cocci’ (Ogston, 1881). These occurred in clumps as opposed to the ‘chain cocci’ previously described within the micrococci family. He went on in 1882 to name this group of isolates ‘Staphylococcus’. This was derived from the Greek word ‘staphyle’ for bunch of grapes (Ogston, 1882). The staphylococci are a family of Gram positive, catalase positive, non-motile, non-spore forming, facultatively anaerobic cocci which range in size from 0.7 – 1.2 µm in diameter. As previously mentioned, these organisms are characterised by their ability to form clusters. They however also occur singly, in pairs or tetrads. Over 30 species of Staphylococcus have been described to date. Of these, the most pathogenic and prominent is S. aureus. This group of isolates derives its species name from the characteristic golden coloured colony. One distinguishing characteristic of this group in relation to other pathogenic Staphylococcus spp of clinical significance is in its ability to produce the coagulase enzyme. This characteristic has led to the general delineation of staphylococci into S. aureus and coagulase negative Staphylococci (CoNS). In addition, S. aureus strains are characterised by their ability to tolerate high salt concentrations (7.5 – 10% w/v NaCl) and ferment mannitol (Lowy, 1998).
**Staphylococcus aureus**

While the initial focus on *S. aureus* was due to its status as the most common bacterium in pus (Power, 1886), it additionally plays a significant role in abscesses. This organism was also associated with both recurrent and chronic infections (West, 1911, Wright, 1904, Hiss and Zinsser, 1909). *S. aureus* is however a member of the normal human flora. This commensal organism commonly found in the anterior nares and other moist areas of the skin, was earlier reported to be carried persistently by 20 – 30% of the population with 60 – 70% of people intermittently carrying this organism (Kluytmans et al., 1997, Lowy, 1998). A more recent survey however puts the carriage figures as 20% for persistent carriers, 30% intermittent and about 50% of the population as non-carriers (Wertheim et al., 2005). Despite this commensal status, *S. aureus* is almost notorious in its remarkable ability to cause a wide array of diseases (Archer, 1998, Casey et al., 2007) capable of affecting almost all organs. These infections range from the often mild skin and soft tissue infection to potentially life threatening conditions such as bacteraemia (Table 1.1). These infections are often endogenous, arising due to a breach either in the skin or mucosal barriers. This breach then permits spread of the organism into areas where it is capable of causing disease (Lowy, 1998). This ability of *S. aureus* is linked with an impressive repertoire of virulence factors, with a direct causal relationship sometimes noted between specific factors and certain disease conditions.
**Table 1.1: Spectrum of diseases associated with *S. aureus*.**

<table>
<thead>
<tr>
<th>Infection Type</th>
<th>Skin and Soft tissue infection</th>
<th>Others (Invasive infection)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin abscess,</td>
<td>Cellulitis,</td>
<td>Bacteraemia,</td>
</tr>
<tr>
<td>Folliculitis,</td>
<td>Impetigo,</td>
<td>Osteomyelitis,</td>
</tr>
<tr>
<td>Furunculosis,</td>
<td>Necrotizing fasciitis,</td>
<td>Pneumonia,</td>
</tr>
<tr>
<td>Pyomyositis,</td>
<td>Scalded skin syndrome,</td>
<td>Septic shock,</td>
</tr>
<tr>
<td>Myositis,</td>
<td>Surgical site infections,</td>
<td>Conjunctivitis,</td>
</tr>
<tr>
<td>Peritonitis</td>
<td></td>
<td>Meningitis,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pyelonephritis,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UTI,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Septic arthritis,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Infective endocarditis,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Necrotizing pneumonia,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Toxic shock syndrome,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Infective endocarditis,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pleural emphyema,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulmonary abscesses,</td>
</tr>
</tbody>
</table>
1.3 Virulence Determinants

*S. aureus* produces numerous cell surface-located, as well as secreted virulence factors, which aid it in virtually all stages of infection. Some of these virulence factors are directly responsible for specific disease-associated features. An example of this is toxic shock syndrome caused by toxic shock syndrome toxin 1 (TSST-1) (Lowy, 1998). Other virulence factors protect the organism from the host immune response, thus enabling colonisation and/or more successful infection. Examples of these include the *S. aureus* microcapsule which prevents phagocytosis (Wertheim et al., 2005), coagulase which activates the formation of fibrin clots (O'Riordan and Lee, 2004) and protein A which inhibits the opsonic capabilities of IgG by binding to the Fc region of this molecule (Archer, 1998). Protein A – IgG binding also prevents complement activation and helps prevent early detection of invading organisms (Lambris et al., 2008). In addition to protein A, several other factors aid in complement evasion. Staphylokinase reacts with plasminogen to generate plasmin, an active serine protease capable of degrading C3 and C5 complement proteins and thereby inhibiting activation of the complement cascade (Rooijakkers et al., 2005b). Extracellular fibrinogen-binding protein (Efb) binds C3b and inhibits opsonisation (Jongerius et al., 2007). The staphylococcal complement inhibitors (SCINs) block complement activation by stabilising the C3 convertase in a state that is non-functional (Rooijakkers et al., 2005a). Other complement evasion molecules include the staphylococcal superantigen-like protein-7 (SSL-7) and the chemotaxis inhibitory protein of *S. aureus* (CHIPS) (Lambris et al., 2008).

Another set of factors, notable for their associated enzymatic capabilities, enhance invasion and spread by breaking down host tissues. Examples of these include the
proteases, hyaluronidase, elastases and lipase (Lowy, 1998, Gordon and Lowy, 2008). *S. aureus* also produces a wide array of toxins. Some of these are named for the system they affect (e.g. staphylococcal enterotoxins, exfoliative toxin) and others the disease condition they create (TSST-1). Other toxins are named based on the specific cell types on which they act on (e.g. leucocidins and alpha haemolysin). In addition, both TSST-1 and a number of staphylococcal enterotoxins are notable for their ability to stimulate a broad repertoire of T cells, earning these entities the label of superantigens (Lowy, 1998, DeLeo et al., 2009).

The sum total of these virulence factors makes *S. aureus* a successful pathogen capable of affecting nearly all human systems. It is pertinent to note however, that no one strain produces all these factors and often, increased virulence of a strain is directly proportional to the number of virulence determinants carried (Blair, 1962, Peacock et al., 2002b).

This strain to strain variation in virulence factors primarily results from genetic exchange between these organisms. Several virulence genes are part of the *S. aureus* core chromosomal genome. These genes are conserved within > 95% of the species and encode virulence factors expressed by most *S. aureus* isolates. Examples of these include protein A, clumping factor, fibrinogen-binding protein A, coagulase, alpha toxin, lipase and superoxide dismutase (Moore and Lindsay, 2001). A number of other virulence genes in *S. aureus* while currently integrated into the chromosome, are extra-chromosomal in origin. These genes have been acquired as a result of horizontal transfer between strains on mobile genetic elements (MGE). For example, the genes encoding staphylokinase, exfoliative toxin A, Panton Valentine leucocidin,
staphylococcal enterotoxin A, P, and K, are all carried on phages (Betley and Mekalanos, 1985, Coleman et al., 1989, Kaneko et al., 1998, Yamaguchi et al., 2000, Kuroda et al., 2001, Goerke et al., 2009). The genes for enterotoxin D and exfoliative toxin B, on the other hand, are carried on plasmids (Bayles and Iandolo, 1989, Yamaguchi et al., 2000).

In addition to phages and plasmids, a third notable group of MGEs are the pathogenicity islands (PAIs). PAIs make a major contribution to a strain’s pathogenicity and are distinct in that other MGEs such as transposons, phages and plasmids may be found within a PAI. PAIs are a subset of genomic islands (GEIs), named due to the presence of gene clusters which encode virulence factors (Hacker and Kaper, 2000). Like other types of GEIs, PAIs are large discrete genomic regions with a different GC content to the rest of the host chromosome. In addition, other criteria used to define a genomic region as a PAI include: 1. Presence of the PAI in the genome of pathogenic strains of a species but general absence in non-pathogenic strains. 2. Carriage of more than one virulence gene. 3. Characteristic insertion at specific regions of the host chromosome associated with the tRNA loci. 4. Size range of 10 to 180 kb. 5. Presence of flanking repeated sequences. 6. Instability, with the occurrence of spontaneous deletions. 7. Carriage of mobility genes (such as integrases or transposases) which aid in the insertion and deletion of the PAI (Knapp et al., 1986, Hacker et al., 1990, Hacker et al., 1997, Blum et al., 1994).

The discrete genomic regions now referred to as PAIs, were first described in *E. coli* 536, a haemolytic uropathogenic strain of *E. coli* (Knapp et al., 1986). This study mapped the *hly* genes which code for haemolysin, to chromosomal inserts. These
inserts were flanked by 16 bp direct repeats, often spontaneously deleted and absent in non-haemolytic strains. The term PAI was subsequently applied to these characteristic ‘chromosomal inserts’ (Hacker et al., 1990) following the observation that deletion of a similar ‘chromosomal insert’ in E. coli 536 and other E. coli strains resulted in the loss of two linked virulence gene clusters.

Several PAIs have been described in S. aureus (Novick, 2003) and are referred to as staphylococcal pathogenicity islands (SaPIs). These are associated with the genes encoding staphylococcal enterotoxins B, C, K and L and the toxic shock syndrome toxin (TSST-1) gene (Johns and Khan, 1988, Lindsay et al., 1998). SaPIs have been described as discrete chromosomal elements ranging in size from 15 – 20 kb, which are found at specific locations and mobilised specifically by certain staphylococcal phages. The prototype SaPI is SaPI1, the first PAI to be described in S. aureus and gram positive organisms as a whole (Lindsay et al., 1998). SaPI1 is a 15.2 kb discrete genomic region containing the genes which encode for TSST-1 (tst) as well as staphylococcal enterotoxins K and L. This MGE is flanked by 17 nucleotide direct repeats. Additionally, mobility in SaPI1 has definitively been shown to be mediated by a generalised transducing phage, 80α.

Of the different MGEs, some were acquired early in the evolutionary history of certain lineages and passed down vertically. Horizontal transfer of MGE and hence virulence factors into susceptible lineages has been the predominant factor in the evolution of virulent clinical strains (Lindsay et al., 2006). Not surprisingly, MGEs also play a second major role in S. aureus – the development of drug resistance.
1.4 Antibiotic Resistance

In addition to the wide array of virulence factors produced, and its ability to cause disease, *S. aureus* is also renowned for its remarkable ability to rapidly develop resistance to every single antimicrobial agent developed against it. With the large scale production of penicillin in 1943 (Anon, 1943) and its subsequent widespread clinical use, a golden age seemed to have arrived with respect to the treatment of *S. aureus* infections. Up to this point *S. aureus* infections had been associated with high mortality rates (Murray, 1934, Skinner and Keefer, 1941). This euphoria was however short lived as penicillin-resistant isolates were detected as early as 1947. One study published during this period, noted an increase in incidence of resistance from 12.5% to 38% in a single year (Barber, 1947, Friedmann, 1948). Resistance was subsequently linked with the production of penicillinase, a β-lactamase enzyme first described in 1944 (Kirby, 1944). β-lactamase hydrolyses the penicillin β-lactam ring leading to inactivation of antibiotic activity. The genes encoding β-lactamases were thought to be extra-chromosomal and found to be carried generally on plasmids (Richmond, 1965b). Occasionally though, these genes were found on the chromosome and thought to be linked with the staphylococcal β-lactamase transposon Tn552 (Novick, 1963, Richmond, 1965a, Novick and Richmond, 1965, Asheshov, 1969, Lyon and Skurray, 1987).

Over the years, new antibiotics were developed to combat *S. aureus*, but this pathogen was found to evolve mechanisms to thwart the actions of each drug at a very alarming rate (Schito, 2006, Lyon and Skurray, 1987). This resulted in widespread prevalence of multi-drug resistant strains, creating a bleak outlook for the prognosis of *S. aureus* infections rather similar to that of the pre-antibiotic era.
1.4.1  Meticillin And Development Of Meticillin Resistance

By the end of the 1950s, penicillin-resistant strains of *S. aureus* which were often also multi-drug resistant (Jantet, 1959, Bradley and Meynell, 1961), had become endemic in hospitals, apparently selected for by the widespread use of antibiotics (Kempe, 1956). Several surveys noted a rising incidence of such strains with incidences of up to 95.9% (Munch-Petersen and Boundy, 1962) and 92% (Koch and Donnell, 1957) reported in association with concomitant increases in mortality rates. Munch-Petersen and Boundy (1962) provided a summary of reports of penicillin resistance over a period of 17 years. This study involved an analysis of 248 articles published between 1942 and 1959. Despite the overall increase in incidence, a wide variation in the reported incidence of penicillin resistance was noted and ranged from 0% - 95.9%. This wide variation was however a function of differences in the study populations. For the studies which reported on isolates obtained from individuals outside of hospital, the maximum reported incidence of penicillin resistance was 60.4%. In contrast, studies involving hospital isolates, reported incidences of up to 95.9%.

This rising incidence appeared to be driven specifically by penicillin resistance despite the multi-antibiotic resistant nature of these isolates. The discovery of meticillin, a beta-lactamase resistant penicillin, in the early 1960s, was therefore considered to herald in a new age in antimicrobial chemotherapy. Right from the start, this semi-synthetic derivative of penicillin (Figure 1.1) showed great promise. Several studies demonstrated the ability of meticillin to successfully clear a range of infections caused by penicillin-resistant *S. aureus*, often with dramatic results (Harding and Thompson,

**Figure 1.1:** Illustration of the chemical formula of penicillin (A) and meticillin (B)

[Link to the chemical formula image]
Resistance of *S. aureus* isolates to meticillin was first detected in 1961, even prior to widespread use of the drug in clinical settings (Jevons, 1961). However at this time, reports of meticillin resistance were few and far between. Furthermore, a lack of emergence of resistance development was reported, even following prolonged therapy (Barber, 1961, Stewart, 1961). Hence, meticillin resistance was not expected to constitute a major problem in clinical practice. In the following decade and a half, the success of meticillin as frontline therapy for infections caused by these penicillin-resistant and often multi-drug resistant *S. aureus* strains, appeared to support this theory. Despite isolated reports of meticillin resistance from all across the globe (Çetin et al., 1962, Borowski et al., 1964, Gravenkemper et al., 1965, Dornbusch et al., 1969, Maniar, 1969, Kayser, 1975), by the mid-1970s, MRSA had not yet been recognised as a major public health issue. Ominously however, an increase in incidence of MRSA had been noted in a number of publications during this period. Turner and Cox reported an increase from 5.4% to 16.6% in Liverpool General Hospital (UK) between 1963 and 1965 (Turner and Cox, 1967). In 1970, Ridley reported a 10% increase in 5 years, which was in contrast to that of strains received for phage typing by the Public Health Laboratory Service, Colindale (UK) where the incidence in 1964 was 0.6%, an increase of 0.4% from that of 1960 (Ridley et al., 1970). A higher incidence of 55.6% was reported in a burns ward in Birmingham UK in 1975, compared to that of 39.8% in 1967 (Ayliffe et al., 1977). Overall, despite the increase in incidence, at this time, MRSA was still not yet deemed a major public health issue (Parker and Jevons, 1964, Skirrow and Rogers, 1971, Bran et al., 1972).

When meticillin resistant strains were detected, they were predominantly found in hospitals, where they were sometimes responsible for severe outbreaks with high
mortality. Several of these strains were termed ‘Epidemic MRSA’ (EMRSA), so named for their ease of spread within hospitals (Benner and Kayser, 1968, Michel and Priem, 1971, Noone and Griffiths, 1971). From the late 1970s, the situation with regards to meticillin resistance changed. A dramatic rise in reports of colonisation and infection by MRSA strains was noted. Reports from Australia mentioned epidemic proportions by 1979, up from 18% noted in 1970 (Gillespie et al., 1984). In New York, researchers noted an increase in MRSA strains submitted for phage typing from 125 between 1977 – 1978 to 2309 in the year 1982 – 1983 (Schaefler et al., 1984). At the University of Mississippi Medical Centre, a rise in incidence of MRSA strains from 11% in 1979 to 50% in 1981 was observed (Markowitz et al., 1983). In Ireland 302 MRSA isolates were detected from superficial site infections in 1980 compared to 20 in 1976. Over the same time period, the number of MRSA strains isolated from deep seated infections increased from 3 to 118 (Cafferkey et al., 1983). Reports similar to these were quite widespread (Giamarellou et al., 1981, Vickery et al., 1986, Boyce et al., 1981, Townsend et al., 1984, Boyce et al., 1983). One common feature of these reports was the recognition of a change in the nature of circulating MRSA strains. These were noted as being distinct from previously characterised strains of that time, by belonging to a different phage group or showing a novel resistance pattern and also by their ability to spread between hospitals. First reports of these EMRSA strains in England were made in 1981 (Law and Gill, 1988) with spread both within and between hospitals noted. These epidemic strains were often responsible for serious infection, were recalcitrant to control measures and accounted for up to half of all nosocomial staphylococcal infections (Duckworth et al., 1988, Law and Gill, 1988, Kerr et al., 1990, Cooke and Marples, 1987).
Over the following decades, MRSA strains became pandemic, constituting a major public health problem worldwide but generally remained hospital-associated. In one study, 74% of spinal cord injury patients were reported to have acquired the organism following hospital admission (Thom et al., 1999). Furthermore, MRSA isolates were notorious for causing outbreaks in critical care units, affecting patients with reduced immunity and often resulting in prolonged length of hospitalization, with high levels of mortality (Reshad et al., 1994, Doebbeling, 1995, Blot and Hoste, 2001, Murphy et al., 2001, Hunt et al., 1988, Martin et al., 1989, Reboli et al., 1989). Rather interestingly, pandemic strains of epidemic MRSA occurring in more than one country, were found to belong to only a handful of lineages and are thought to have evolved independently (Enright et al., 2002).

1.4.2 Mechanism of Meticillin Resistance

1.4.2.1 mecA Mediated Meticillin Resistance

From the first reports, meticillin resistance was found to be both intrinsic and heterogenous. Intrinsic in the sense that resistance was not associated with penicillinase production which would result in drug inactivation, and heterogenous, with respect to the fact that expression of resistance varied between cells in a culture grown at 37°C. Only a small proportion of these cells exhibited the resistant phenotype (Barber, 1961). Research has since linked meticillin resistance with the presence of a unique/novel class of penicillin-binding-protein (PBP-2A) which is predominant in MRSA but absent in MSSA. PBP-2A exhibits a much reduced affinity for beta lactam antibiotics (Brown and Reynolds, 1980, Hartman and Tomasz, 1984, Hartman and Tomasz, 1981). The 2130 bp mecA gene which encodes for this
transpeptidase, together with its regulatory genes mecI and mecRI make up the mec gene complex.

Though located on the chromosome (Sjostrom et al., 1975, Kuhl et al., 1978), the mec gene complex is carried by a mobile genetic element, the staphylococcal cassette chromosome mec (SCCmec). At present, the exact origin of this mobile genetic element has not been determined. The detection of non-mecA-containing SCC elements in several species of CoNS however led researchers to conclude that this mobile genetic element most likely originated from the CoNS (Mongkolrattanothai et al., 2004, Katayama et al., 2003, Kuroda et al., 2005, Takeuchi et al., 2005, Hanssen and Ericson Sollid, 2006). Isolates carrying these SCC elements, presumably then acquired the mec gene complex to form SCCmec. The mecA gene itself is however thought to have evolutionary links with the coagulase-negative Staphylococcus sciuri. A study of thirteen different Staphylococcus spp of animal origin, found that a mecA gene homologue with 80% nucleotide identity was present in all 134 strains S. sciuri examined (Couto et al., 1996).

The SCCmec element was first described in 1998 (Ito and Hiramatsu, 1998) and was found to consist of three main components, the mec and ccr gene complexes and the regions separating these which are termed the junkyard or joining regions. Variations occur in all these regions leading to different SCCmec types as defined by the combination of mec and ccr (Table 1.2). By 2009, eight different SCCmec types had been described (Figure 1.2) (Ito et al., 1999, Katayama et al., 2000, Ito et al., 2001, Ma et al., 2002, Ito et al., 2004, Oliveira et al., 2006, Berglund et al., 2008a, Zhang et al., 2009, Okuma et al., 2002).
Table 1.2: Composition of the eight SCC\textit{mec} types described by 2009.

<table>
<thead>
<tr>
<th>SCC\textit{mec} Type</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>type 1 \textit{ccr}/class B \textit{mec}</td>
</tr>
<tr>
<td>Type II</td>
<td>type 2 \textit{ccr}/class A \textit{mec}</td>
</tr>
<tr>
<td>Type III</td>
<td>type 3 \textit{ccr}/class A \textit{mec}</td>
</tr>
<tr>
<td>Type IV</td>
<td>type 2 \textit{ccr}/class B \textit{mec}</td>
</tr>
<tr>
<td>Type V</td>
<td>type 5 \textit{ccr}/class C \textit{mec}</td>
</tr>
<tr>
<td>Type VI</td>
<td>type 4 \textit{ccr}/class B \textit{mec}</td>
</tr>
<tr>
<td>Type VII</td>
<td>type 5 \textit{ccr}/class C\textsubscript{1} \textit{mec}</td>
</tr>
<tr>
<td>Type VIII</td>
<td>type 4 \textit{ccr}/class A \textit{mec}</td>
</tr>
</tbody>
</table>

Adapted from \url{www.sccmec.org/Pages/SCC_TypesEN.html}
Figure 1.2: Schematic representation of variation in the eight SCCmec types described by 2009 showing the different composition and arrangement of mec (in pink) and ccr gene complexes (in blue).

The isolate identities are indicated in brackets.

The figure was reproduced from: A report by the International working group on the classification of staphylococcal cassette chromosome elements (Anon, 2009).
1.4.2.2 Non-\textit{mecA} Mediated Resistance

In addition to the \textit{mecA} mediated meticillin resistance exhibited by the majority of MRSA isolates, several studies have reported cases of non-\textit{mecA} mediated meticillin resistance. These studies described a distinct group of isolates which exhibited low level resistance (oxacillin MIC of 2 – 8 µg/ml), lacked the \textit{mec} gene complex and furthermore, were characterised by uniform resistance with no detectable numbers of highly resistant cells. These strains were termed, borderline resistant \textit{S. aureus} (BORSA). This resistance was found to be mediated by two different mechanisms depending on the \textit{S. aureus} strain. One group of researchers (McDougal and Thornsberry, 1986) reported an association between staphylococcal \textit{β}-lactamase production and BORSA. In this study, the presence of \textit{β}-lactamase was found to have an effect on the MIC of oxacillin for \textit{β}-lactamase positive BORSA strains but not for \textit{β}-lactamase positive MRSA strains. For the \textit{β}-lactamase positive BORSA strains, a significant reduction in MIC was noted when a \textit{β}-lactamase inhibitor (clavulanic acid or sulbactam) was added during susceptibility testing (2 – 0.27 µg/ml). Thus indicating a \textit{β}-lactamase mediated meticillin borderline resistance. In contrast, the addition of \textit{β}-lactamase inhibitor during susceptibility testing of \textit{β}-lactamase positive MRSA isolates, had no effect on oxacillin MIC. This \textit{β}-lactamase mediated meticillin borderline resistance was confirmed by several other researchers (McMurray et al., 1990, Montanari et al., 1990, Petersson et al., 1999).

This phenomenon was found to occur predominantly in phage group 94/96 (also known as phage group V) isolates, due to the presence of a 17.2 kb \textit{β}-lactamase-expressing plasmid (pBW15) (McMurray et al., 1990, Zierdt et al., 1992). The
findings of one study further suggested that the borderline meticillin resistance is a phenomenon specific only to isolates carrying the pBW15 plasmid. This suggestion was based on the observation that hyper-production of β-lactamases by several other strains which lacked the pBW15 plasmid failed to result in borderline resistance (Barg et al., 1991).

An earlier study characterising BORSA (Tomasz et al., 1989) had identified a second possible mechanism of non-meca mediated meticillin resistance. In this study, β-lactamase negative BORSA were found to produce ‘normal’ PBP which however differed in their penicillin binding capabilities and the levels produced. PBP-1 and PBP-2 were found to exhibit lower binding capabilities while PBP-4 was produced in elevated levels. No mention was however made of the phage grouping for the isolates in this study.

At present though, no comprehensive data exists for the prevalence of these BORSA and researchers are divided on the clinical significance of these strains especially with respect to treatment failures (Swenson et al., 2001, Petersson et al., 1999, Massanari et al., 1988, Kernodle et al., 1990).

1.5 Community-acquired MRSA

Prior to the mid-1990s, the majority of MRSA strains worldwide were reported circulating in hospitals, driven by the combination of high antibiotic use and patient’s impaired immunity. These strains were notorious for persistence and spread and often linked with a high degree of morbidity and mortality (Cooke and Marples, 1987, Law
and Gill, 1988, Thompson et al., 1982, Martin et al., 1989, Peacock et al., 1980). By the late 1990s, a change in this epidemiological trend of MRSA was observed. Increasing reports were made of MRSA infections in community settings in patients with no known hospital links. Sporadic reports of community-acquired MRSA (CA-MRSA) had been made over the years, but these strains had generally been described among intravenous drug users or in long term care facilities, such as nursing homes (Saravolatz et al., 1982, Markowitz et al., 1983, Storch et al., 1987). Newly emerging strains of CA-MRSA were however not linked to these populations but rather affected healthy individuals with no known risk for MRSA infections (Herold et al., 1998, Price et al., 1998, CDC, 1999, L'Heriteau et al., 1999, Gwynne-Jones and Stott, 1999, Berman et al., 1993). Several studies reported outbreaks among members of sports teams and in healthy teenagers (Lindenmayer et al., 1998, Stacey et al., 1998). A study carried out at the University of Chicago Children’s Hospital to ascertain CA-MRSA prevalence noted an increase from 10 per 100,000 admissions between 1988 and 1990 to 259 per 100,000 admissions between 1993 and 1995. In this study, isolates were defined as CA if obtained within 72 h of hospitalization (Herold et al., 1998). Similarly, Purcell and Fergie (2002) reported an increase in prevalence of CA-MRSA from 2.9% of all S. aureus isolates in 1990 to 40.3% in 2001 at Driscoll Children’s Hospital Texas. While the majority of infections described involved skin and soft tissue (Prevost et al., 1995a, Frank et al., 1999, Gorak et al., 1999, Bukharie et al., 2001), reports were made of more severe infections such as endocarditis and necrotising pneumonia (Mallory et al., 1997, Gillet et al., 2002, Villar et al., 1999), often in previously healthy and young individuals.
Concurrently, some studies reported a surge in colonisation of the general public. These studies noted CA-MRSA colonisation highs ranging from 3.3% to 13.2% (Hidron et al., 2005, Vlack et al., 2006, Huang et al., 2007, Lo et al., 2007, Lo et al., 2008, Lu et al., 2005, Creech et al., 2005, Wang et al., 2009, Sdougkos et al., 2008). This was in contrast to studies prior to 2003 which reported CA-MRSA colonisation highs ranging from 1.5% to 3% (Abudu et al., 2001, Suggs et al., 1999, Kenner et al., 2003, Jernigan et al., 2003).

By 2005, CA-MRSA had been linked with more severe outcomes and significant fatalities (Peleg and Munckhof, 2004, Tseng et al., 2005, Francis et al., 2005, Miller et al., 2005, Miyashita et al., 2002). At this time, these strains were also reported to be gradually spreading into healthcare facilities (Klevenet al., 2006, Kourbatova et al., 2005, Saiman et al., 2003).

1.5.1 Characteristics of CA-MRSA strains

No standardized criteria for defining CA-MRSA strains existed and therefore, some variations consequently occurred between researchers (Ispahani et al., 1987, Moreno et al., 1995, Naimi et al., 2001, Herold et al., 1998). The key points in CA-MRSA definition however focused on the timing of isolation and healthcare links. There were significant differences in the details. In one study, an isolate was termed CA if detected within 24 h of hospitalization of the patient (Naimi et al., 2001). Other studies reported strains as CA if MRSA was isolated within either 48 h (Ispahani et al., 1987, Moreno et al., 1995) or 72 h (Herold et al., 1998) of hospitalization. Other additional factors sometimes used in defining S. aureus isolates as CA included:
history of prior hospital contact or hospitalisation, antibiotic history, history of injection drug use or presence of an indwelling medical device (Hughes et al., 1976).

One significant feature not addressed in these definitions was the source of the infecting strains. While these so called ‘CA-MRSA’ could potentially have been HA-MRSA spread to communities via discharged patients, healthcare visitors or healthcare workers, initial reports revealed an unusual antimicrobial susceptibility profile of CA-MRSA. Unlike the generally multi-resistant HA-MRSA, the CA-MRSA isolates demonstrated susceptibility to most antimicrobial agents apart from the beta-lactam antibiotics. Additionally, these strains were found to belong to different, often more diverse lineages from those previously associated with HA-MRSA, as described by both PFGE patterns and multilocus sequence typing (MLST) (Purcell and Fergie, 2002, Herold et al., 1998, Abi-Hanna et al., 2000, Block and Carmichael, 1978, Okuma et al., 2002).

The final factor supporting the hypothesis of independent evolution of these CA-MRSA strains by SCCmec acquisition by MSSA strains in the community, rather than a spread of HA-MRSA, was the variation in SCCmec elements found among CA isolates. Unlike their hospital counterparts, CA-MRSA carried a then novel SCCmec element, the type IV SCCmec element. These type IV elements were notably smaller than the previously described SCCmec types, with a size range of 21 – 25 kb and lacking in non β-lactam resistance genes (Figure 1.2). This resulted in the unique resistance phenotype that was initially noted and a higher fitness level than the hospital-acquired strains. CA-MRSA appeared to have evolved independently in specific locations, varying in clonal distribution with a representation among more
diverse genetic backgrounds (Daum et al., 2002, Okuma et al., 2002, Cookson, 2000). In addition, CA-MRSA strains were found to multiply faster \textit{in vitro}. Okuma et al. (2002) reported a significantly lower mean doubling time of $28.79 \pm 7.09$ min for CA-MRSA in contrast to the mean doubling time of $38.81 \pm 7.01$ min observed for the HA-MRSA strains.

1.6 Panton-Valentine leucocidin

1.6.1 Link to CA-MRSA

Following the establishment of MRSA in the community, several studies explored the unique characteristics of these strains, especially with respect to virulence, and identified a common factor among CA-MRSA strains. This factor was the PVL toxin (Vandenesch et al., 2003, Said-Salim et al., 2003, Dufour et al., 2002). The genes encoding this toxin were found to be highly represented within this population. Up to 100% prevalence was observed among CA-MRSA strains as opposed to a prevalence of only 2% observed in the general \textit{S. aureus} population (Kilic et al., 2006, Kilic et al., 2008, Makgotho et al., 2009, Prevost et al., 1995a). A higher prevalence of PVL within distinct populations of \textit{S. aureus} was initially noted in a 1995 French publication (Prevost et al., 1995a). This study looked at two different \textit{S. aureus} populations. The first group of isolates comprised 309 consecutive clinical strains, 106 of which were from tracheal aspirations and sputa. Within this population, only 1.6% (5) of isolates carried the genes encoding for PVL. These isolates were subsequently designated PVL-positive. In sharp contrast to this low prevalence, 16.8% of a second group of \textit{S. aureus} isolates were PVL-positive. This second group comprised 346 isolates originating specifically from either blood cultures, cutaneous infections or
nasal carriage. The majority (246, 71.1%) of these isolates however were from cutaneous infections.

A similar higher prevalence of PVL-positive isolates was noted in a 1999 study (Lina et al., 1999). In this study which analysed 172 selected community acquired isolates from the French National Reference Center for Staphylococci (Lyon, France) strain collection, 37.2% were PVL-positive. The epidemiological link of PVL with CA-MRSA was subsequently further strengthened over the following years. Several publications dealing with a wide range of sample sizes from various countries reported prevalence’s of PVL-positive CA-MRSA ranging from 8% – 100%, most of which were between the range of 69.4% and 100% (Diep et al., 2004, Liassine et al., 2004, Vandenesch et al., 2003, Frazee et al., 2005, Yamasaki et al., 2005, Wannet et al., 2004, Faria et al., 2005, David et al., 2008, Larsen et al., 2009b, Cheung et al., 2008, Huang et al., 2008, Vourli et al., 2009). Notable among these studies was the 2003 report by Vandenesch and colleagues (Vandenesch et al., 2003). This study analysed 117 CA-MRSA isolates from six countries (United States, France, Switzerland, Australia, New Zealand and Western Samoa). 100% of these CA-MRSA isolates were PVL-positive. In contrast, PVL encoding genes were not detected in the 57 HA-MRSA strains analysed simultaneously. This low prevalence of PVL-positive isolates in either the general S. aureus population or HA-MRSA specifically, was also reported in other studies (Vandenesch et al., 2003, Holmes et al., 2005, Huang et al., 2008, Liu et al., 2009, Vindel et al., 2009, Kilic et al., 2006, Kilic et al., 2008) and ranged from 0% – 20%, with the majority below 2%.
Initially, strains carrying the genes encoding PVL were more commonly associated with certain infections such as primary skin infections and community-acquired pneumonia rather than deep seated infections or urinary tract infections (Lina et al., 1999, Prevost et al., 1995a, Cribier et al., 1992, Finck-Barbancon et al., 1991). Reports of PVL-associated necrotising pneumonia were not as widespread as those for skin and soft tissue infection, but a rise in incidence of necrotising pneumonia occurred concurrently with an increase in prevalence of strains carrying the genes encoding PVL. The predilection of this often fatal infection for young apparently healthy, immune-competent individuals, and the high level of mortality associated with necrotizing pneumonia however made it a serious public health concern (Le Thomas et al., 2001, Gillet et al., 2002, Osterlund et al., 2002, van der Flier et al., 2003, Hyvernat et al., 2007, Gillet et al., 2008).

Subsequently, strains of CA-MRSA emerged worldwide causing diverse types of infections and are epidemic in the USA due to the spread of the single USA300 clone. These have spread into healthcare facilities amid reports of increased antibiotic resistance (Bocchini et al., 2006, Enany et al., 2007, Ramos et al., 2009, Sdougkos et al., 2007, Tsai et al., 2008, Diep and Otto, 2008, Moran et al., 2006, Linde et al., 2005, Bratu et al., 2005, Gonzalez et al., 2006, Tristan et al., 2007).

In the UK, though the incidence of CA-MRSA is relatively low and reports infrequent (Elston and Barlow, 2009), the Health Protection Agency (HPA) has noted a steady rise in the number of PVL-positive MRSA. These PVL-positive MRSA however, account for less than 50% of the total number of reported cases of infection by PVL-positive *S. aureus* (Table 1.3).
Table 1.3: Rising incidence of PVL-positive *S. aureus* isolates in the UK.

<table>
<thead>
<tr>
<th>Year</th>
<th>MSSA PVL</th>
<th>MRSA PVL</th>
<th>Total PVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>107</td>
<td>117</td>
<td>224</td>
</tr>
<tr>
<td>2006</td>
<td>337</td>
<td>159</td>
<td>496</td>
</tr>
<tr>
<td>2007</td>
<td>729</td>
<td>477</td>
<td>1206</td>
</tr>
<tr>
<td>2008</td>
<td>1013</td>
<td>724</td>
<td>1738</td>
</tr>
</tbody>
</table>

Data from HPA Colindale 2009.

Data reflects PVL-positive isolates referred to the HPA based on clinical suspicion (recurrent boils, abscesses, necrotising skin and soft tissue infection, community-acquired necrotising pneumonia) and a unique gentamicin/trimethoprim antibiotic resistant profile (HPA, 2008, Boakes et al., 2011b).
These HPA figures were however not generated following systemic surveillance. Rather, the figures represent PVL-positive isolates identified from strains referred to the Staphylococcal Reference Unit (SRU) for PVL toxin profiling based on clinical suspicion as described in the HPA guidance on diagnosis of PVL-positive cases of *S. aureus* infection (HPA, 2008). This data could therefore reflect an increase in case detection and not necessarily provide a true representation of PVL prevalence in England and Wales. In contrast to reports from the United States where most PVL-positive strains are MRSA USA300, an analysis of UK data indicates that infection by PVL-positive MSSA is more common than that by PVL-positive MRSA (Ellington et al., 2007, Hussain et al., 2007, Mushtaq et al., 2008, Thomas et al., 2009, Cunnington et al., 2009, Shallcross et al., 2009).

Over the years several reports have been made of the association of PVL-positive MSSA with infection worldwide (Nolte et al., 2005, Adler et al., 2006, Schleucher et al., 2008, Kalka-Moll et al., 2008). A study by Gillet and colleagues in 2007 noted a higher incidence of PVL-MSSA as opposed to PVL-MRSA as a cause of necrotising pneumonia (Gillet et al., 2007). These CA-MSSA strains differ from their CA-MRSA counterparts in having a more diverse genetic background (Miller et al., 2007). They however exhibit similar clinical and epidemiological characteristics (Golding et al., 2010, Bocher et al., 2008, Miller et al., 2007), resulting in indistinguishable infections (Murray et al., 2004). PVL-positive strains of MSSA have similarly been associated with more severe infection (Schefold et al., 2007, Mongkolrattanothai et al., 2003, Perbet et al., 2010, Swaminathan et al., 2006, Lorenz et al., 2007, Ramos et al., 2009). Though the majority of research worldwide generally focuses on CA-MRSA carrying
the PVL encoding genes, several reports of a higher prevalence of PVL-positive MSSA strains rather than PVL-positive MRSA strains (Severin et al., 2008, Baranovich et al., 2009, Sila et al., 2009), points at a looming threat that has most probably been overlooked.

1.6.2 PVL Origins

The PVL toxin was first described in 1932 (Panton and Valentine, 1932). This bi-component toxin was initially separated into what he termed the slow and fast fractions based on the speed of elution during ion exchange chromatography (Woodin, 1959). These two components, LukF and LukS, were initially reported to have molecular weights of 32 kDa and 38 kDa (Woodin, 1959) and 38 kDa and 32 kDa (Finck-Barbancon et al., 1991) respectively. LukF and LukS have however now been shown to be proteins of 34 kDa and 32 kDa respectively (Prevost et al., 1995b), which act synergistically to produce the active PVL toxin.

Based on its mode of action, the PVL toxin is associated primarily with two families of toxins. The first of which is the synergohymenotrophic toxins. This family first described in 1993, is so named because members contain two components acting synergistically as membrane damaging toxins (Supersac et al., 1993). Members of this family include the gamma haemolysin and LukE-LukD leucocidins. Together with the staphylococcal alpha toxin, they also form part of a larger family, the β-barrel pore forming toxins (PFT). These are secreted as water soluble entities that bind and assemble on target cell membranes to form pores (Menestrina et al., 2003).
Several differences occur between the PVL toxin and other members of these two families of toxins. While production of gamma haemolysin has been noted in approximately 99% of all *S. aureus* strains, only 2% produce PVL. PVL also differs in being toxic only for leukocytes rather than for erythrocytes (Kaneko and Kamio, 2004, Prevost et al., 1995a). This toxicity is further host specific. A thousand times more toxin is required to result in toxicity in mouse and monkey polymorphonuclear cells (PMNs) than compared to human and rabbit cells (Loffler et al., 2010). Initial studies showed that PVL interacted with the cell membrane of human and rabbit leukocytes resulting in various cellular abnormalities. These abnormalities include a loss of motility, withdrawal of pseudopodia, disappearance of granules, alterations of cell membrane permeability and cell swelling. Mouse, sheep and guinea-pig leukocytes however, appeared resistant to these effects (Gladstone and Van Heyningen, 1957, Woodin and Wieneke, 1966, Grojec and Jeljaszewicz, 1981). Furthermore, the effect of PVL on leukocytes was found to be limited to PMNs, monocytes and macrophages. Lack of toxicity for lymphocytes is presumably due to the observed lack of a LukS binding site on these cells (Gauduchon et al., 2001). Unlike the alpha toxin, which forms heptameric pores, PVL forms octameric pores composed of 4 units of both the LukF and LukS subunits (Miles et al., 2002).

Studies by Genestier and colleagues in 2005 pointed at a variation in pore size created by the PVL toxin. This variation was dependent on the concentration of toxin present. Higher concentrations of toxin were shown to result in the formation of larger pores generally causing necrosis. At low toxin concentrations, cells tended to undergo apoptosis. The ability of the PVL toxin to induce apoptosis was thought to result from the action of the toxin on the mitochondrion following entry of PVL molecules into
the cells when present at low concentration (Genestier et al., 2005). This toxin has also been noted to induce the liberation of inflammatory mediators such as histamine, cytokines and chemokines (Hensler et al., 1994, Tseng et al., 2009), as well as exhibit potent dermonecrotic characteristics (Ward and Turner, 1980).

1.6.3 Genetics of PVL

The two PVL subunits are coded for by a pair of adjacent co-transcribed genes, *lukF* and *lukS* (Prevost et al., 1995b), which together make up the ≈1.9 kb *lukSF-PV* locus (Figure 1.3).

**Figure 1.3: Schematic representation of the *lukSF-PV* locus**

The *lukSF-PV* locus is composed of the *lukS* and *lukF* genes in the order shown in the figure. Numbering corresponds to nucleotide positioning of the *lukS* and *lukF* genes with relation to the *lukSF-PV* locus.
The lukF and lukS gene sequences are highly conserved within PVL-positive S. aureus strains, with only 12 major single nucleotide polymorphisms (SNPs) described to date (Takano et al., 2008b). Most of the mutations in the lukSF-PV locus are synonymous, not resulting in an amino acid change. Of the non-synonymous mutations, the SNPs at nucleotides 181 and 470 which result in a change from phenylalanine to valine and tyrosine to phenylalanine, respectively (Takano et al., 2008b), occur only in minor variants. More focus has been on the major non-synonymous mutation found at nucleotide 527. This causes a histidine to arginine amino acid change at amino acid position 176, reported to change the predicted molecular size of the LukS protein subunit from 35,294 Da to 35,313 Da (Kaneko et al., 1998, Kaneko et al., 1997, Narita et al., 2001, Baba et al., 2002, Diep et al., 2006b, Ma et al., 2006, Ma et al., 2008, Takano et al., 2008b, Berglund et al., 2008b, Otter et al., 2010, Wolter et al., 2007). The possibility of this widespread H176R mutation resulting in functional implications for the toxin was hypothesized based on molecular modelling (O'Hara et al., 2008). However, direct studies on toxicity in two independent studies established that the H176R mutation does not result in significant variations in leucotoxicity and pore forming capabilities of the resulting toxin (Berglund et al., 2008b, Besseyre des Horts et al., 2010).

Interestingly, the sequence variations present in the genes encoding the LukF and LukS proteins appear to correlate with evolutionary origins of a specific strain of S. aureus. Members of a specific multilocus ST have generally been reported to have the same lukSF-PV sequence variations (Table 1.4).
Table 1.4: Relationship between *lukSF-PV* sequence variants and evolutionary origins

<table>
<thead>
<tr>
<th>ST</th>
<th>33</th>
<th>105</th>
<th>181</th>
<th>216</th>
<th>345</th>
<th>470</th>
<th>527</th>
<th>663</th>
<th>1396</th>
<th>1729</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>8/1518</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>22/59</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>30/50</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>80</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>88</td>
<td>G</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
<td>C</td>
<td>NA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>Phage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>øPVL108/øSLT</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>ø2958PVL</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>øPVL</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
</tbody>
</table>

Compiled from (Takano et al., 2008b, Berglund et al., 2008b, Otter et al., 2010, Wolter et al., 2007)

NA: Data Not Available

ST denotes multilocus sequence types associated with the various *lukSF-PV* sequence variants

The three last rows of the table show the *lukSF-PV* sequence variations associated with some PVL phages
1.6.4 **PVL phages**

The genes encoding PVL were subsequently shown to be extra chromosomal in origin, carried by specific phages, of which at least six have been described in the literature (øSLT, øPVL, ø108PVL, ø2958PVL, øSa2USA and øSa2mw) (Kaneko et al., 1998, Kaneko et al., 1997, Narita et al., 2001, Baba et al., 2002, Diep et al., 2006b, Ma et al., 2006, Ma et al., 2008). The six temperate bacteriophages described were similar in size ranging from 41,401 bp to 45,924 bp, but composed of two different distinct phage head morphologies, icosahedral and elongated. While only 15% identity was noted between the DNA sequence of øSLT and øPVL, the PVL genes were highly conserved within all phages with a 99.8% identity observed.

1.6.5 **PVL Controversy**

With an apparently firmly established epidemiological link between PVL and CA-MRSA, the focus of PVL research subsequently shifted to demonstrating a role for PVL as a major virulence factor in CA-MRSA infections. Several studies were carried out between 2006 and 2009 to answer this question (Labandeira-Rey et al., 2007, Brown et al., 2009, Cremieux et al., 2009, Voyich et al., 2006, Bubeck Wardenburg et al., 2007a, Bubeck Wardenburg et al., 2008, Diep et al., 2008, Montgomery and Daum, 2009). These studies looked for possible effects of the PVL toxin in CA-MRSA infections, using different animal and infection models, and both isogenic PVL-positive and PVL-negative strains, but resulted in varied and sometimes contradictory conclusions.
On one hand, results from several studies pointed at a role for PVL as a virulence factor in osteomyelitis, pneumonia and skin and soft tissue infections (Labandeira-Rey et al., 2007, Brown et al., 2009, Cremieux et al., 2009). These studies reported that infection with PVL-positive isolates resulted in a higher degree of morbidity, stronger inflammatory response and increased levels of mortality at 24 h when compared with infection with PVL-negative strains in different animal infection models examples of which include models of pneumonia model and skin infection model in mice and a rabbit osteomyelitis model.

On the other hand, a 2008 study by Diep and colleagues using a mouse bacteraemia model (Diep et al., 2008), reported findings which could perhaps indicate a role for the PVL toxin early in infection. Animals infected with PVL-positive strains had a higher bacterial density in the kidneys at 24 h and 48 h post infection than when compared to infection with isogenic PVL-negative strains. However, by 72 h a similar bacterial density was observed in the kidneys of animals infected with either the PVL-positive or the PVL-negative isogenic strains.

A third group of researchers however failed to clearly demonstrate a role for PVL in virulence (Voyich et al., 2006, Bubeck Wardenburg et al., 2007a, Bubeck Wardenburg et al., 2008, Diep et al., 2008, Montgomery and Daum, 2009). In these studies, no difference was observed between PVL-positive and their PVL-negative isogenic S. aureus strains with respect to morbidity, mortality, PMN lysis or plasma membrane damage. One study, involving a mouse pneumonia model, even noted a higher degree of mortality following infection with PVL-negative strains (Bubeck Wardenburg et al., 2008).
At first glance, though the notion that similar research could produce widely differing results provided a source of puzzlement, closer scrutiny of study design provided clues for the discrepancies between results of these experiments. These study designs were not identical but rather, differed on several fronts. Various animal and infection models were represented within the different studies. Examples of these include a mouse pneumonia model, a mouse skin and soft tissue model, a rabbit osteomyelitis model, a rat pneumonia model and a rabbit disseminated infection model. More studies though, made use of the mouse model of infection. Mice PMNs have however been previously reported to be relatively insensitive to PVL in comparison to human or rabbit PMNs (Gladstone and Van Heyningen, 1957, Woodin and Wieneke, 1966, Grojec and Jeljaszewicz, 1981). Hence, questions have been raised on the suitability of the mouse model in analysing the contribution of PVL to disease virulence and the possible effect this could have on experiment outcome (Schwartzman et al., 2007, Loughman et al., 2009). Furthermore, reports that in contrast to BALB/cAnNHsd mice, no significant mortality occurred in BALB/c mice following intranasal inoculation with $5 \times 10^7$ cfu of either PVL-positive or PVL-negative isogenic LAC strains Bubeck Wardenburg et al. (2008) pointed at subtle variations in susceptibility to PVL even between mice strains.

In addition to the variation in animal infection models, the different studies also further varied in the type of *S. aureus* strain used. Examples of the different *S. aureus* strains utilized in these studies include two laboratory *S. aureus* strains (RN6390 and BK9918) and three standard *S. aureus* strains (Newman, LAC and MW2). Further variations were noted in the mode of inoculation, method of sample preparation for
analysis of morbidity, criteria for defining morbidity, growth media employed and inoculum size.

Based on these variations in study design several issues were raised. These included (1) the effect of using a S. aureus strain containing a plasmid expressing high levels of PVL (Labandeira-Rey et al., 2007) on result validity, (2) the accuracy of the parameters used to define infection in the animal models (with analysis of myositis beneath wound infections shown to provide better discrimination than a cursory look at just the superficial wound size), (3) the importance of timing in assessing weight loss and (4) the procedure for determining lung inflammation (Brown et al., 2009, Bubeck Wardenburg et al., 2008, Bubeck Wardenburg et al., 2007a).

The importance of some of these key factors in disease outcome was eventually addressed systematically in a 2009 study using a mouse model (Tseng et al., 2009). This study involved the use of a larger inoculum \(10^9\) cfu than previously described \(10^7\) cfu in the Voyich and Bubeck studies (Voyich et al., 2006, Bubeck Wardenburg et al., 2008). Use of a \(10^9\) cfu inoculum clearly showed differences between muscle lesions caused by PVL-positive strains compared with PVL-negative strains in CD1 and BALB/c mice genotypes. No differences were however noted between skin lesions produced following infection with either PVL-positive or PVL-negative strains. Furthermore, this link between PVL-positive strains and severity of muscle lesions was also shown to depend on mouse strain. Unlike the results observed with the CD1 and BALB/c mice genotypes, for C57/B6 and SKH1 mice, no differences were observed between muscle lesions caused by PVL-positive strains compared with PVL-negative strains. This correlated with the level of cytokine induction. In the CD1
and BALB/c mice, the level of macrophage inflammatory protein-2 (MIP-2) and keratinocyte-derived chemokine (KC) were significantly higher following infection with PVL-positive strains compared with PVL-negative strains (≈42,000 pg/ml versus ≈10,000 pg/ml and ≈25,000 pg/ml versus ≈12,000 pg/ml for MIP-2 and KC respectively in CD1 mice; ≈39,000 pg/ml versus ≈27,000 pg/ml and ≈42,000 pg/ml versus ≈20,000 pg/ml for MIP-2 and KC respectively in BALB/c mice). In contrast, infection of C57/B6 and SKH1 mice with PVL-positive strains, did not elicit significantly higher level of chemokine (≈41,000 pg/ml versus ≈39,000 pg/ml and ≈38,000 pg/ml versus ≈31,000 pg/ml for MIP-2 and KC respectively in C57/B6 mice; ≈26,000 pg/ml versus ≈27,000 pg/ml and ≈21,000 pg/ml versus ≈24,000 pg/ml for MIP-2 and KC respectively in SKH1 mice). Higher levels of cytokine induction was also found in younger mice and correlated with increased degrees of PVL injury.

When this information is linked with the observation that culture medium composition greatly affects the expression of the genes encoding the LukF and LukS PVL subunits (Diep et al., 2008), with CCY resulting in higher expression of the PVL toxin than TSB, one would tend to agree with the suggestion put forward on the limitations of current in vitro models used for studying PVL (Loughman et al., 2009) and the need for a standardized protocol.

1.6.6 PVL Expression Studies

One other factor that could significantly affect the outcome of infection studies using PVL-positive S. aureus strains is the actual level of toxin expressed in vivo. Several of the initial studies attempting to link PVL and virulence had completely failed to address this factor. Approaches to the investigation of PVL expression have varied
both in the samples analysed and the method used to quantify the level of PVL expression. Different studies explored PVL expression both directly in skin abscess samples (Loughman et al., 2009, Badiou et al., 2008) and in vitro, in culture supernatants (Hamilton et al., 2007, Said-Salim et al., 2005). Additionally, one study quantified anti-PVL antibody titre in serum in a bid to relate this to the level of toxin present (Croze et al., 2009). Though Bronner and colleagues had previously expressed concern about the difficulty of leucotoxin quantification by ELISA, due to cross reactivity with other class S or class F components (Bronner et al., 2000), advances in research meant this issue has been resolved. Hence PVL expression in these studies was explored with one of two techniques: Enzyme linked immunosobent assay (ELISA) to detect the toxin itself, or reverse transcriptase RT-PCR which directly ascertained lukS/lukF mRNA transcript levels.

In all these studies, PVL expression was detected in all strains harbouring the encoding genes, albeit with differing levels noted. Also, PVL expression in tissues was reported to be up-regulated when compared to in vitro levels of expression by the same strains (Loughman et al., 2009). However, separate studies found no correlation between in vitro levels of PVL expression and disease severity (Hamilton et al., 2007, Said-Salim et al., 2005). Additionally, PVL expression was found to be effected by the presence of neutrophils as noted by a 2009 study analysing lukS mRNA transcript levels in vitro (Loughman et al., 2009). This study reported an up-regulation of PVL expression when clinical strains of S. aureus were grown in media containing neutrophils as opposed to when the same strains were grown in media alone. These findings were in agreement with a later study (Tseng et al., 2009) which showed increased pathogenic effect of PVL-positive S. aureus in young mice with a more
aggressive immune response and hence a higher neutrophil level. In addition to the neutrophil level, several other factors have been shown to affect the expression of the \textit{lukF} and \textit{lukS} genes. Some of these include antibiotic concentration (Dumitrescu et al., 2008a) and the virulence global regulators \textit{agr} and \textit{sae} (Bronner et al., 2000, Wirtz et al., 2009). A reduction in \textit{sae} expression was reported to result in decreased PVL production. In addition, the temporal expression of the PVL toxin has also recently been linked to the life cycle of the encoding phage and was also influenced by the genetic background of the \textit{S. aureus} host (Wirtz et al., 2009). These various reports serve to highlight current gaps in the knowledge of factors affecting PVL expression, both \textit{in vivo} and \textit{in vitro}.

Though progress is being made in relation to understanding the role of PVL in virulence and pathogenesis, a lot of unknowns still exist. In the meantime, the prevalence of PVL-positive CA-MRSA infections and colonisations, continue to increase. Several studies have proposed that several factors other than PVL may be responsible for the apparently enhanced virulence of community-acquired strains of \textit{S. aureus}. These include alpha haemolysin and the phenol soluble modulins (PSMs).

1.6.7 Alpha Haemolysin/Toxin

Unlike PVL which came into recent prominence due to the rise in CA-MRSA, the role of the alpha toxin (HLA) in virulence had been established as far back as 1991. As evidenced by studies from 1940 aimed at enhancing the production of HLA (Casman, 1940), and research on its composition (Bernheimer and Schwartz, 1963), the potential importance of this toxin was clearly recognised and a debate on its role in pathogenesis begun as early as 1963 (Taubler et al., 1963).
This pore forming toxin, which is a prototype of the beta barrel toxin family and the first pore forming bacterial cytolysin identified, is the most widely studied of the staphylococcal toxins. By 1991, several researchers had extensively studied this molecule and provided most of the background knowledge summarized below (Bhakdi and Tranum-Jensen, 1991).

The alpha toxin, initially thought to be as large as 44 kDa, is a 33.2 kDa monomer which binds to a specific target membrane, assembles to form a mushroom shaped heptamer and subsequently undergoes conformational changes to form a 1 – 2 nm functional pore. This toxin is active on a wide variety of cell types such as lymphocytes, erythrocytes, platelets, endothelial cells, macrophages, fibroblasts and keratinocytes. These may differ in susceptibility. Human platelets and rabbit erythrocytes have been noted to exhibit a higher sensitivity than human erythrocytes. These are significantly resistant, requiring up to 400-fold more toxin to induce lysis (Bernheimer and Schwartz, 1963). Based on the concentration of toxin present, binding of the alpha toxin monomer may occur either via specific or nonspecific interactions. Pore formation often leads to membrane injury and a bidirectional flow of ions, water and molecules of low molecular weight such as ATP and ultimately, cell lysis (Bhakdi and Tranum-Jensen, 1991, Walev et al., 1993). Secondary effects such as alteration of activities of immune cells, stimulation of pro-inflammatory mediators (Dragneva et al., 2001) and induction of programmed cell death (apoptosis) have also been described (Menzies and Kourteva, 2000, Jonas et al., 1994).
Over the years, alpha toxin has been confirmed as a staphylococcal virulence factor in a wide variety of infection types (Bramley et al., 1989, Patel et al., 1987, Callegan et al., 1994, Kielian et al., 2001). In addition, its toxoid has been demonstrated to offer protection against infection (Hume et al., 2000). It was only recently though that this staphylococcal virulence factor began to be linked specifically with CA-MRSA, and suspected of playing a role in the virulence of these strains. Specifically in relation to *S. aureus* pneumonia, *S. aureus* LAC *hla* mutants failed to produce disease in a mouse host and immunisation with anti-*hla* antibodies was shown to protect from infection (Bubeck Wardenburg et al., 2007a, Bubeck Wardenburg et al., 2007b, Ragle and Bubeck Wardenburg, 2009).

The gene encoding the alpha toxin in contrast to the PVL encoding genes, is found on the chromosome of about 99% of *S. aureus* strains. Expression of this gene may vary under different growth conditions, and is influenced by the presence of specific antibiotics such as β-lactams and clindamycin as well as quinupristin/dalfopristin (Ohlsen et al., 1998, Koszczol et al., 2006, Herbert et al., 2001, Ohlsen et al., 1997). Furthermore, expression is controlled by regulatory systems such as the accessory gene regulator (*agr*), staphylococcal accessory gene regulator (*sar*), extracellular protein regulator (*xpr*), *S. aureus* exoprotein expression (*sae*), repressor of toxin (*rot*) and multiple gene regulator (*mgr*) (Novick et al., 1993, Hart et al., 1993, Cheung et al., 1992, Giraudo et al., 1997, McNamara et al., 2000, Luong et al., 2003). The specific role of alpha haemolysin in disease caused by community-acquired strains of *S. aureus* has however not yet been verified.
1.6.8 Phenol Soluble Modulins (PSMs)

Phenol soluble modulins (PSMs) belong to a relatively novel class of virulence factors ‘the modulins’. These modulins stimulate cytokine production in host cells (Henderson et al., 1996). PSMs were specifically named due to their solubility in phenol and were initially described in *S. epidermidis*. These factors exhibit pro-inflammatory characteristics similar to other modulins previously described, thereby indicating a possible role in pathogenesis (Mehlin et al., 1999, Liles et al., 2001).

In 2007, a group of researchers described a novel α-type PSM (PSMα) related to the PSM found in *S. epidermidis* (Wang et al., 2007). While the genes encoding PSMα were found to be carried on the genome of all sequenced *S. aureus* strains, the level of *in vitro* production of the peptides differed. Wang and colleagues noted a higher expression of PSMα in CA-MRSA compared to HA-MRSA. Furthermore, this study reported that PSMα deletion mutants resulted in reduced mortality in a mouse bacteraemia model and exhibited a reduced capacity to induce the lysis of human PMNs. These findings led these researchers to propose a role for PSM as the major factor associated with the leucocytic ability of CA-MRSA strains and hence their virulence. This role has however been questioned by a more recent report which observes that, though CA-MRSA isolates produce increased amounts of PSM, the concentrations of PSM produced are not sufficient to cause cell lysis (Hongo et al., 2009). The study by Hongo and colleagues did however note an enhancement of lysis of human PMNs by PVL in the presence of PSMα.

This research into the role of PSMs in CA-MRSA is however still in its preliminary stages.
1.7 *S. aureus* typing

To better understand the role of PVL in virulence of community-acquired strains of *S. aureus*, as well as the effect of sequence variations in the *lukF* and *lukS* genes on epidemiology and spread of these strains, the ability to accurately type, identify related isolates and adequately discriminate between variants has been key.

Strain characterisation is important in understanding the epidemiology (spread, relatedness and significance) of a particular group of organisms, as well as helping in the prevention and control of outbreaks. Over the years, several typing systems have been used in strain characterisation. These typing systems have evolved from phenotypic systems to the more recent genotypic typing systems and differ in their discriminatory power, reproducibility and relevance to specific isolate (Jones et al., 1974, Kashbur et al., 1974, Cookson et al., 1992, Old et al., 1980, Scherer and Stevens, 1987, Colding et al., 1999, Singh et al., 2006). The most important *S. aureus* typing systems will now be considered in detail.

1.7.1 PFGE

With respect to *S. aureus* typing, pulsed field gel electrophoresis (PFGE) is the gold standard. PFGE which was first described by Schwartz and Cantor in 1984 is a modification of the standard agarose gel electrophoresis method which enables the resolution of much larger DNA molecules (up to 2,000 kb), by alternating the orientation of the electrical currents applied during electrophoresis (Schwartz and
Cantor, 1984). This method which was originally used in eukaryotic studies, was first applied as a *S. aureus* typing method in 1991 (Prevost et al., 1991). PFGE involves the separation of DNA fragments generated following a restriction enzyme digestion, to give a distinct restriction pattern for each isolate. Though several restriction enzymes were initially tested, digestion of *S. aureus* genomic DNA prior to PFGE is routinely carried out using the *SmaI* enzyme. The *SmaI* enzyme was found to produce the best outcome, resulting in 15 – 20 fragments, ranging in size from 30 to 1,500 kb. Other enzymes tested, resulted in fragments which were either too many or too few (Prevost et al., 1991, Ichiyama et al., 1991, Struelens et al., 1992, Schlichting et al., 1993). Following PFGE, isolates can then be typed as same or different by visual comparison of the restriction patterns produced.

PFGE was found to have very high discriminatory powers and mainly proving useful in short term epidemiology studies, with its ability to distinguish between isolates which were evolutionary related but epidemiologically independent (Arbeit et al., 1990). Several studies have since confirmed the superiority of PFGE to other typing techniques including phenotypic analysis, plasmid profiling, phage typing, ribotyping, multilocus enzyme electrophoresis (MLEE) and random amplification of polymorphic DNA (RAPD) (Prevost et al., 1992, Saulnier et al., 1993). PFGE however has the disadvantage of being technically demanding, labour intensive and often difficult to interpret, which may result in poor inter-laboratory reproducibility. Hence alternative approaches to typing *S. aureus* strains have been investigated.
With DNA sequencing technology becoming more commonplace, sequence based typing systems have been developed in recent years. Two such systems are multilocus sequence typing (MLST) and spa typing.

1.7.2 **Multilocus Sequence Typing (MLST)**

The multilocus sequence typing method which was first described in 1998 (Maiden et al., 1998), is derived from the cumbersome MLEE. MLEE uses allele specific variation in the electrophoretic mobilities of several housekeeping enzymes to generate strain fingerprints. Like PFGE, MLEE was a standard method in eukaryotic population genetics before its adaptation to bacterial genetics. The MLEE method detects enzyme variants based on differences in enzyme mobility, and is based on the principle that this variation relates directly to allelic variations in the encoding genes. As these genotypic differences were not expected to always result in enzyme variants, this technique involved the analysis of a large number of loci (15 – 25 enzymes) to provide sufficient discrimination. While MLEE enabled strain characterisation providing global epidemiology the technique was often laborious, technically demanding and subjective, making inter-laboratory comparison difficult (Selander et al., 1986, Musser et al., 1990, Musser and Kapur, 1992).

Rather than looking at variations in electrophoretic mobility of enzymes, MLST focuses directly on variations in the DNA sequences of internal fragments (≈450 bp) of housekeeping genes. Thereby allowing the study of fewer loci and resulting in the classification of isolates into sequence types (STs). As these genes evolve slowly over time, MLST gives a picture of the evolutionary history and global epidemiology for
individual strains. Currently, the MLST protocol has been developed for 29 microorganisms (http://www.mlst.net/databases/default.asp, assessed 11th September 2012) including many of the major pathogens. The protocol developed for typing strains of *S. aureus*, found that variations in just seven key housekeeping genes are enough to class isolates into groups of related ancestry (Enright et al., 2000). The housekeeping genes used, and their corresponding gene products are *arcC* (carbamate kinase), *aroE* (shikimate dehydrogenase), *glpF* (glycerol kinase), *gmk* (guanylate kinase), *pta* (phosphate acetyltransferase), *tpi* (triosephosphate isomerase) and *yqiL* (acetyl coenzyme A acetyltransferase). In MLST, following amplification and sequencing of fragments of the seven genes, each isolate is assigned an allele number at each of the loci. The information from each individual locus is combined to form what is termed an allelic profile which then corresponds to a sequence type (ST). For example, an isolate belonging to sequence type (ST) 1 has an allele number of one for each locus and hence an allelic profile of 1-1-1-1-1-1-1.

The fact that the MLST technique monitors several genes which evolve slowly over time, eliminates the possibility of a chance occurrence when two isolates have the same allelic profile. This same characteristic is also a disadvantage of the technique as studying a number of genes makes the technique cumbersome and rather expensive to perform. One of the major advantages of the MLST method is that it is a nucleotide sequence based method with the results available online. As such, it is highly portable, allowing both reproducibility and inter-laboratory comparison, quite unlike PFGE. With several studies clearly demonstrating similar discriminatory abilities as PFGE, (Peacock et al., 2002a, Melles et al., 2007, Grundmann et al., 2002), MLST is now one of the major systems used for typing of *S. aureus* isolates.
1.7.3  *spa Typing*

A second typing technique routinely used in studying the epidemiology, prevalence and spread of *S. aureus*, is the *spa* typing method. Use of variations in the *spa* gene which encodes the staphylococcal protein A (Spa) as an epidemiologic tool was proposed initially in 1994 (Frenay et al., 1994). Protein A, one of the major surface proteins of *S. aureus* is a cell wall protein covalently linked to the peptidoglycan. Spa consists of an IgG binding region and a polymorphic X region (*Figure 1.4*) made up of a variable number of 24 bp repeats (Guss et al., 1984). The variations in this region often arise from deletions, duplications and point mutations in the repeats (Kahl et al., 2005). The 1994 Frenay study found a correlation between the number of 24 bp repeats in the X region of protein A and the epidemic status of the strain. Most epidemic strains were found to have > 7 repeats while non-epidemic strains contained 7 or less repeats. The actual *spa* typing technique involves the amplification and sequencing of the polymorphic X region using primers located in the well conserved flanking regions (Frenay et al., 1996). Spa types are then assigned based on the sequence, number and arrangement of repeats found.
Figure 1.4: Illustration of the spa gene showing the polymorphic X region ($X_r$) used in spa typing and the positioning of the primer sequences used for both amplification and sequencing.

Reproduced from: Shopsin et al. (1999)

The numbered arrows indicate the 5’ start location of the forward and reverse primers on a *S. aureus* (GenBank accession No: J01786) forward strand. The spa gene segments are indicated by the lettered boxes (not to scale): S: Signal Sequence; A – D: Immunoglobulin G-binding regions; E: Region homologous to A – D; $X_c$: Cell wall attachment sequence.
Prior to 2003 and the development of the Ridom StaphType software (http://spaserver.ridom.de/) (Harmsen et al., 2003), several programs had been used for the identification and assignment of new repeats. These programs, unlike the Ridom StaphType, described spa types using alpha numeric codes and differed between researchers, often leading to non-uniform assignment of spa types (Shopsin et al., 1999, Tang et al., 2000). With the introduction of the open access internet-based Ridom StaphType software this limitation has been overcome and portability and uniformity of the technique has been established.

Further validation of this technique as a molecular tool has deemed it comparable to other current techniques with a discriminatory power lower than PFGE but higher than MLST (Shopsin et al., 1999). Thus, spa typing has proved quite useful in the investigation of local outbreaks especially but has also been shown to be valuable in the analysis of the long term evolutionary history of S. aureus strains (Shopsin et al., 1999, Khandavilli et al., 2009, Mellmann et al., 2008). In addition to speed, other advantages include ease of use, interpretation and inter-laboratory comparison. Currently, MLST is a more widely used technique in relation to S. aureus epidemiology but, due to its cost advantage, spa typing is becoming more commonly used with researchers applying spa typing to larger panels of isolates and MLST to only a subset of such strains (Ellington et al., 2009, Alp et al., 2009, Witte et al., 2007).

So far over six thousand (6,000) spa types have been described containing a total of 364 unique repeats with the number of repeats in each spa type ranging from 2 to 16
Though a few exceptions occur, most spa types are unique to an MLST clonal complex (Basset et al., 2009, Deurenberg et al., 2007, Cookson et al., 2007) making both techniques comparable.

1.7.4 **High Resolution Melt Analysis**

Current typing methods though quite discriminatory, are however still quite cumbersome, expensive and time consuming. High resolution melt (HRM) analysis is an extension to the traditional melting curve capabilities of real-time PCR systems and is one recently introduced technology which may have useful application as a rapid diagnostic tool, providing useful typing information. The first reports of HRM in 2003, used labelled primers in a closed-tube system to identify sequence variants of the β-globin gene without any additional manipulations prior to performing the melt analysis (Gundry et al., 2003). The use of labelled primers however meant that variations in the sequence could only be detected if they occurred within the same melt domain as the labelled primer. A subsequent advancement to this method published in the same year overcame this limitation. This method involved the use of a novel dye (LCGreen) that unlike SYBR Green, bound to DNA but did not inhibit the PCR at high concentrations (Gundry et al., 2003, Wittwer et al., 2003).

Since then, HRM has been widely applied in both gene scanning – often leading to disease detection – (Krypuy et al., 2006, Takano et al., 2008a, White et al., 2007, Slinger et al., 2007) and genotyping of different isolates based on variations in specific genes. Examples include the detection and identification of several clinically important bacterial species by variations in their 16S rRNA gene (Cheng et al., 2006),
the differentiation of influenza A virus subtypes based on the viral matrix gene (Lin et al., 2008), and *Mycoplasma pneumoniae* typing (Schwartz et al., 2009).

HRM combined improved dye chemistry and machine resolution with novel software technologies to detect differences in amplicon melt curves with a high degree of sensitivity and specificity. Originally, dyes used in the detection of double stranded DNA such as SYBR Green I, were non-saturating and showed an inhibitory effect on DNA amplification at high concentrations. The development of saturating ds-DNA binding dyes was pivotal in HRM development. These dyes unlike SYBR Green I are non-inhibitory and may therefore be used at saturating concentrations, thus enabling binding to each base pair of the target amplicon. This property also possibly prevents dye redistribution during the melting process and results in improved sensitivities and accuracies (Wittwer et al., 2003). LCGreen was the first saturating dye to be described and others developed over the years include SYTO9, LCGreen Plus, EvaGreens and resolight dye (Wittwer et al., 2003, Monis et al., 2005, Vossen et al., 2009). HRM is currently available on several platforms, examples of which include the ABI 7300, Eppendorf Mastercycler RealPlex4S, Idaho Technology LightScanner, Roche LightCycler480 and Corbett Rotor-Gene 6500HRM. In comparison to the real time platforms, these instruments have an improved ability to measure amplicon melt behaviour due to an increase in temperature precision and improved data acquisition. The accompanying software is a third factor contributing to HRM sensitivity. This software allows the generation of a difference curve from the melt curve, usually following a temperature shift to correct for subtle temperature differences across the plate, and a normalization process which corrects for variations in starting template concentration (Herrmann et al., 2007, Herrmann et al., 2006).
HRM is capable of detecting a single base pair difference in an amplicon of up to 1,000 bp in length (Reed and Wittwer, 2004). This capability of HRM however depends in part on amplicon length. Several studies have reported a higher degree of sensitivity for shorter length products (Reed and Wittwer, 2004, Gundry et al., 2003, Liew et al., 2004, Krypuy et al., 2006, Herrmann et al., 2006). Best results were obtained when analysing HRM products of less than 300 bp in length. In addition to amplicon length, the SNP class affects the outcome of HRM analysis. SNP classes describe the grouping of the six possible binary combinations of base mutations (where a G – T mutation is equivalent to a T – G mutation) based on the type of base change (Liew et al., 2004). Class 1 SNPs represent a C – T or G – A transition mutation (i.e. purine to purine change or pyrimidine to pyrimidine change). On the other hand, class 2 SNPs represent the C – A or G – T transversion mutation (i.e. purine to pyrimidine change). Class 3 SNPs specifically refer to a G – C mutation and class 4, a T – A mutation. While all SNPs can potentially be detected, class 3 and 4 SNPs produce relatively small Tm differences (0.25°C) compared to the class 1 and 2 SNPs (0.8 – 1.4°C) and hence often result in more problematic genotyping.

1.8 Typing approaches used to study evolution of S. aureus

One major contribution made by the several typing systems utilized in the study of S. aureus, was to the understanding of the evolution of this group of isolates, particularly the meticillin resistant strains. While the initial phenotypic typing systems simply enabled the classification of isolates as either same, different or possibly related, the early molecular based systems MLEE and PFGE went a step further.
MLEE defined genetic distance between two groups of isolates using the average linkage method described by Selander and colleagues (Selander et al., 1986). In this method, isolates which differed at all loci had a genetic distance of 1 and identical isolates a distance of 0. When combined with the assumption for PFGE that one SNP in chromosomal DNA could create a three fragment variation in restriction profile (Bannerman et al., 1995), both methods recognized that the relationship of isolates with a profile similarity of 85% or more (corresponding to a genetic distance of 0.20) was such that this high degree of similarity could not have occurred by chance. Hence such strains represent subtypes of the same clone with a common ancestor.

For MLEE and PFGE which type isolates based on 15 to 20 variables, this could simply be interpreted thus: isolates with a maximum allelic variation of 3 were considered clonal but were classed as different if the number of mismatches exceeded this number (Selander et al., 1986, Musser and Kapur, 1992, Struelens et al., 1992, Bannerman et al., 1995, Dominguez et al., 1994). Tenover and colleagues went on to further propose that since 2 independent genetic events could possibly result in a 4 – 6 band difference, isolates differing at less than 7 bands were probably of the same genetic lineage but only less closely related (Tenover et al., 1995). These guidelines were however only applied in an outbreak setting.

The development of MLST (Enright et al., 2000) enabled major breakthroughs in the understanding of how these bacterial clones evolved. For this technique in which isolates were typed based on genetic variations at 7 housekeeping genes (which by their nature evolve relatively slowly over time), isolates identical at all loci were thought to have evolved from a common ancestor and classed as a sequence type (ST).
Those differing at one locus were however considered closely related and grouped together to form a clonal complex (Feng et al., 2008). One study went on to demonstrate how well the evolutionary events as defined by MLST agreed with that previously proposed by Tenover and colleagues based on PFGE band variation (Enright et al., 2000). Isolates belonging to the same sequence type were found to differ at 4 bands or less while members of the same clonal complex generally had a maximum of 7 band differences. Conversely, highly dissimilar STs possessing variations in 6 of the 7 MLST loci were found to have at least a 20 band difference.

Perhaps the major success of MLST as a technique was in its ability to provide a framework to work out the process whereby various lineages evolved from each other. The putative ancestral ST for each clonal complex was usually the predominant member of the group. This had long been defined as the genotype which varied from most other members of the CC at only a single locus. Hence, subsequent members of each clonal complex were thought to have evolved from these putative ancestral STs (Enright et al., 2002). This hypothesis was validated by the findings of Feil and colleagues published in 2003 (Feil et al., 2003). In this study, single locus variants (SLVs) were shown to possess only a single nucleotide difference from the putative ancestral ST. Furthermore, the resulting variant locus was generally not represented among any other ST but unique to the specific SLV, clearly demonstrating that these STs had evolved directly from the putative ancestral genotype via point mutations rather than recombination. At that time, two groups of isolates (ST34 and ST40) were found to appear as exceptions to this model of evolution. Both groups contained alleles not unique to their clonal complexes and hence could only have arisen via recombination between two different clonal complexes (Feil et al., 2003).
1.8.1 eBURST

One tool which proved useful in studying the evolution of *S. aureus*, was the eBURST (based upon related sequence types) algorithm described in 2004 (Feil et al., 2004). This algorithm was a new implementation of the previously described BURST program (Enright et al., 2002) which was designed primarily for the analysis of MLST data. The principle of the eBURST algorithm is based on the simplest model which explains the emergence of clonal complexes. This model proposes that the emergence of a clonal complex results following an increase in the frequency of a ‘founding genotype’ within a population. This increase, results in genetic diversification of the founding genotype such that descendants which differ from this ‘founding genotype’ at one of seven MLST loci accumulate, resulting in a clonal complex. The eBURST algorithm has three major functions. First, eBURST divides the STs in the MLST database into mutually exclusive groups of related genotypes based on a user defined parameter. This defines members of an eBURST group as STs with a minimum number of identical alleles as at least one other member of the group. The minimum number of identical alleles may be user defined, but more commonly a default setting is used. The default setting defines an eBURST group as STs which are identical at 6 or 7 of the MLST alleles as at least one other member of the group. Based on this definition, STs which differ at two or more alleles from every other ST in the population cannot be assigned a group and are referred to as singletons. Within the eBURST groups, the algorithm then further identifies a ST which serves as the putative ancestral genotype or primary founder from which the other STs diverged. The primary founder is defined as the ST with the largest number of SLVs within the
Importantly, this process does not depend on the number of isolates in each ST and hence, is independent of sampling bias.

Furthermore, the level of confidence in the accuracy of the predicted primary founder is represented by a bootstrap value. These values represent the percentage of times a specific sequence type is predicted as the primary founder following a bootstrap resampling procedure. This need for statistical confidence in the predicted primary founder is important because of the third function of eBURST. This algorithm predicts and displays the most likely pattern of evolutionary descent of the isolates in the clonal complex from this primary founder. This information is presented as a radial diagram whereby the primary founder is linked to its SLVs, which then link to double locus variants (DLVs) of the primary founder and so on.

Despite the current widespread use of this algorithm, it must be noted that the pattern of evolutionary descent described by eBURST is inferred and does not take into account any additional data, be it phenotypic, genotypic or epidemiological. An example of a possible misrepresentation of evolutionary relationship by the eBURST algorithm was reported in the initial eBURST publication (Feil et al., 2004). ST239 was described as a primary founder of a group containing ST8, on the basis of ST239 possessing one SLV more than ST8. This grouping however, contrasted with a previous study (Enright et al., 2002). The study by Enright and colleagues, assigned ST8 as the primary founder of this group of isolates based on the absence of meticillin sensitive strains in ST239 and the observation that ST239 carried a type III SCCmec element found only in an MRSA ST8 strain within this grouping. Feil and colleagues had however reported a similarly high bootstrap value for ST239 and ST8 (70% and
66% respectively). This finding highlighted a possible issue in assigning ST239 as primary founder and emphasized the importance of taking the bootstrap value into consideration when interpreting the data generated by eBURST.

1.8.2 MRSA Evolution

The genotypic typing systems also proved quite invaluable in understanding the evolution of MRSA, especially in deciding between the single and multi–clone theories of MRSA evolution. Prior to 2002, several studies had put forward evidence in support of either different MRSA clones emerging from a single MSSA clone which acquired the SCCmec element once (the single clone theory), or the multiple acquisition of SCCmec elements by the different S. aureus lineages (multi clone theory) (Lacey and Grinsted, 1973, Carles-Nurit et al., 1992, Musser and Kapur, 1992, Kreiswirth et al., 1993, Crisostomo et al., 2001). The extensive study by Enright and colleagues in 2002, however, showed the multi clone theory to be the more likely theory (Enright et al., 2002). This study, which analysed over 900 isolates, noted the presence of both MSSA and MRSA within the same ST. This finding thereby pointed at the development of MRSA following the acquisition of a SCCmec element by meticillin sensitive strains. In addition to this, the detection of multiple SCCmec types in strains belonging to a single ST pointed at a case of multiple acquisitions. An exception to this model of MRSA evolution was however described with ST259 and ST247. These two groups of isolates lack meticillin sensitive members but rather are composed entirely of meticillin resistant strains. Hence ST259 and ST247 are thought to have evolved from point mutations in MRSA variants of their parent clones.
1.9 Restriction-Modification Systems

One factor which may come into play in the evolution of *S. aureus* strains is the bacterial restriction-modification (RM) system. RM systems were initially described in the early 1950s (Luria and Human, 1952, Maio and Zahler, 1958, Bertani and Weigle, 1953) and are mainly found in bacteria but also present among other prokaryotes (Boyer, 1971, Patterson and Pauling, 1985, Wilson, 1991). RM systems are thought to serve as some sort of primitive bacterial immune system whereby they create a barrier that restricts the uptake of foreign DNA. These systems are thought to have evolved primarily to protect bacteria from viral infection. Bacterial RM systems are so named due to their mode of action. They act mainly by degrading foreign DNA fragments following the detection of a specific sequence, while a modification of the hosts own DNA by methylation protects it from restriction. In some cases though, restriction only occurs on modified foreign DNA (Bickle and Kruger, 1993).

Several RM systems exist and differ in their composition and exact mechanism of action. RM systems may contain up to three distinct subunits, the R, M and S subunits (Wilson and Murray, 1991, Wilson, 1991, Bickle and Kruger, 1993, Boyer, 1971). The S or specificity subunit recognises a specific sequence to be modified or restricted, the M or modification subunit methylates the host DNA and the R or restriction subunit facilitates the cleavage of unmethylated DNA. The type I RM system is the most complex (Wilson and Murray, 1991) and has all three subunits present. In this system, the recognition site/sequences are not symmetrical but rather made up of two elements ranging from three to five nucleotides in length separated by a non-specific region ‘the spacer’. This is often six to eight nucleotides long, as seen in the case of *E. coli EcoBI* which recognises the sequence TGA(N₈)TGCT (Dryden et
al., 2001). This recognition site however differs from the cleavage site, with restriction enzymes acting at a significant distance away from it. On the other hand, the type II system is often described as the simplest. This is more commonly found in nature and serves as the source of quite a number of the restriction enzymes employed in biological research. In this case both recognition and cleavage occur at the same sequence. The sequences are specific and cleavage takes place within the sequence in a symmetric manner (Pingoud et al., 2005). Other RM types include the type IIIs which have a non-symmetric recognition sequence and an endonuclease that acts a distance from this. In addition to this, the rarely occurring type III and type IV systems have an asymmetric and uninterrupted recognition sequence (Janulaitis et al., 1992). The type IV differs simply from the type III in that it does not require ATP for its activity.

Several restriction modification systems have been described in strains of S. aureus and while these predominantly belong to the type II class, these systems are not widespread and are often carried on mobile genetic elements (Sussenbach et al., 1976, Iordanescu and Surdeanu, 1976, Sussenbach et al., 1978, Seeber et al., 1990, Godany et al., 2004, Dempsey et al., 2005). Recently however, a chromosomally located type I RM system (SauI) has been described (Waldron and Lindsay, 2006). This SauI system is found to be widely distributed among strains of S. aureus and includes a single restriction gene (hsdR), two copies of the modification gene (hsdM and hsdM1) and two copies of the sequence specificity gene (hsdS1 and hsdS2). Three of these hsd genes (hsdR, hsdM and hsdM1), exhibited a high degree of homology (>95%). Variations in both copies of the specificity gene (hsdS and hsdS1) however, appear linked with S. aureus lineage and hence may have contributed to the development and maintenance of these lineages.
1.10 Conclusion

Since the first reports of CA-MRSA infections in the mid 1990s, the threat it poses has not diminished but rather increased. These strains are now routinely isolated in health care settings with the CA-MRSA USA300 strain reported as the most frequent cause of hospital-acquired MRSA infections in the USA (Seybold et al., 2006). Taken together with the increase in multi-drug resistance observed in these strains (Boyle-Vavra et al., 2005, Takizawa et al., 2005), the very definitions of CA-MRSA are being challenged, blurring the lines between community-acquired and hospital-acquired isolates. The need therefore to define and understand factors which play a role in both the virulence and pathogenesis of these isolates remains ever pressing. Though recent reports confirming a role of PVL in CA-MRSA infections (particularly necrotising pneumonia) have brought the scientific community a step forward in relation to these strains, a lot of questions are still left unanswered.

1.11 Aims and Objectives

This first aim of the present study was to explore the genetic diversity of a group of local PVL-MSSA clinical isolates in order to contribute to the limited current data and provide insight into the evolution and emergence of PVL-MRSA isolates. In addition, as current typing systems are cumbersome and time consuming, the present study was also aimed at the development of a rapid high resolution melt (HRM) typing system for the characterisation of PVL-isolates. Thirdly, an enzyme-linked immunosorbent assay (ELISA) system was developed for the detection and quantification of both PVL
and alpha haemolysin which could be used as a research tool and potentially in clinical diagnosis to provide prompt diagnosis and hence improve patient care.
Chapter Two

2 General Materials and Methods

2.1 Media

Unless otherwise stated all growth media were purchased from either Fisher Scientific (Loughborough, UK), Oxoid (Basingstoke UK) or Sigma-Aldrich (Dorset UK), prepared according to manufacturer’s instructions and sterilised by autoclaving at 121°C at 15 psi for 20 min.

The media used in the present study include Luria Bertani (LB) broth, LB agar, Brain Heart Infusion (BHI) broth, Trypticase Soy Broth (TSB) and CYGP (Casamino acids-yeast extract-glycerophosphate) broth. The composition of these media is as indicated in Appendix 1.

100 µg/ml of ampicillin was added to the media where required to select for the growth of ampicillin-resistant cultures.

2.2 Chemicals

All chemicals used in the present study were obtained from either Sigma-Aldrich (Dorset, UK) or BD Biosciences (Oxford, UK), unless otherwise stated.
2.3 PVL-positive test isolates.

The PVL-positive *S. aureus* used in the present study consisted of two groups, both of which contained MSSA and MRSA isolates. The first group contained recent clinical isolates recovered between November 2008 and May 2009 from Nottingham University Hospitals (NUH) NHS Trust and confirmed as PVL positive by the Health Protection Agency (HPA) Staphylococcus Reference Unit (SRU) (Table 2.1). The second group contained all of the PCR confirmed, PVL-positive strains available in the University of Nottingham *S. aureus* strain collection. These had been obtained from two sources which are indicated in Table 2.2. All of the isolates were stored frozen at -80°C in trypticase soy broth (TSB) containing 10% v/v glycerol.
Table 2.1: NUH NHS Trust PVL-positive *S. aureus* clinical isolates

<table>
<thead>
<tr>
<th>S/No</th>
<th>Strain ID</th>
<th>Meticillin Status</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS1</td>
<td>MSSA</td>
<td>NUH&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>2.</td>
<td>TS2</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
<tr>
<td>3.</td>
<td>TS5</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>4.</td>
<td>TS6</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>5.</td>
<td>TS7</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>6.</td>
<td>TS8</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>7.</td>
<td>TS9</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>8.</td>
<td>TS12</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>9.</td>
<td>TS13</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>10.</td>
<td>TS14</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>11.</td>
<td>TS15</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>12.</td>
<td>TS16</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>13.</td>
<td>TS17</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>14.</td>
<td>TS18</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>15.</td>
<td>TS19</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>16.</td>
<td>TS20</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>17.</td>
<td>TS21</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>18.</td>
<td>TS23</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>19.</td>
<td>TS24</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>20.</td>
<td>TS25</td>
<td>MSSA</td>
<td>NUH</td>
</tr>
<tr>
<td>21.</td>
<td>TS26</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
<tr>
<td>22.</td>
<td>TS27</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
<tr>
<td>23.</td>
<td>TS28</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
<tr>
<td>24.</td>
<td>TS29</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
<tr>
<td>25.</td>
<td>TS30</td>
<td>MRSA</td>
<td>NUH</td>
</tr>
</tbody>
</table>

<sup>a</sup>NUH: Nottingham University Hospital.

Isolates were recovered between November 2008 and May 2009 at NUH and confirmed as PVL positive by the Health Protection Agency (HPA) *Staphylococcus Reference Unit* (SRU)
Table 2.2: PVL-positive *S. aureus* isolates from the University of Nottingham *S. aureus* strain collection.

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID</th>
<th>Meticillin Status</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>NRS123</td>
<td>MRSA</td>
<td>NARSA(^a)</td>
</tr>
<tr>
<td>2.</td>
<td>NRS157</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>3.</td>
<td>NRS158</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>4.</td>
<td>NRS162</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>5.</td>
<td>NRS185</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>6.</td>
<td>NRS192</td>
<td>MRSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>7.</td>
<td>NRS194</td>
<td>MRSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>8.</td>
<td>NRS227</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>9.</td>
<td>NRS229</td>
<td>MSSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>10.</td>
<td>NRS248</td>
<td>MRSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>11.</td>
<td>NRS255</td>
<td>MRSA</td>
<td>NARSA</td>
</tr>
<tr>
<td>12.</td>
<td>RSS289</td>
<td>MSSA</td>
<td>RS(^b)</td>
</tr>
<tr>
<td>13.</td>
<td>RSS290</td>
<td>MSSA</td>
<td>RS</td>
</tr>
</tbody>
</table>

\(^a\)NARSA strains had previously been purchased from the Network on Antimicrobial Resistance in *Staphylococcus aureus* strain collection

\(^b\)RS: Richard Spence strain collection comprised of clinical strains isolated from the Nottingham University Hospitals between January 2003 and October 2005 (Spence, 2007).
2.4 Growth of *Staphylococcus aureus* and *Escherichia coli*

Strains used in the present study were propagated at 37°C aerobically, unless otherwise stated. Growth on solid media was routinely carried out statically by overnight incubation, while liquid cultures were grown overnight in a G25 orbital shaker incubator (New Brunswick Scientific Co Inc. Edison, New Jersey USA) with shaking at 200 rpm.

2.5 Genomic DNA Extraction

DNA extraction from *S. aureus* cultures was carried out using the DNeasy® blood and tissue kit (Qiagen, Crawley UK) according to manufacturer’s instructions with some modifications. Firstly, bacterial cells were harvested from 3 ml of overnight culture by centrifugation at 13,000 rpm for 2 min using a benchtop centrifuge (Pico, Biofuge). Cells were subsequently lysed by resuspending in 180 µl enzymatic lysis buffer (20 mM Tris [pH 8.0], 2 mM EDTA and 1.2% v/v Triton X-100) supplemented with 5 µl of 5 mg/ml lysostaphin in water and 5 µl of 1 mg/ml RNase A in water and incubated for 30 min at 37°C. This was followed by addition of 25 µl of 600 mAU/ml of proteinase K and 200 µl of buffer AL (both supplied with the Qiagen kit) and a 30 min incubation at 56°C. Following this, 200 µl of 100% ethanol was added to the sample and the whole mixture applied to a DNeasy spin column, from which it was purified by a series of washes and centrifugation as described in the manufacturer’s instructions. Finally, the DNA was eluted from the column using 200 µl of sterile distilled water and stored at -20°C.
2.6 Plasmid DNA Extraction from *E. coli*

For plasmid extraction, 3 ml of overnight culture was harvested by centrifugation at 13,000 rpm for 1 min and extraction carried out using the Qiaprep miniprep kit (Qiagen, Crawley UK) according to the manufacturer’s instructions. Cell pellets were resuspended in 250 µl of Buffer P1. Following this, 250 µl of Buffer P2 and 350 µl of Buffer N3 were added to complete the lysis process and denature genomic DNA. Cell debris and precipitated genomic DNA were then removed by centrifugation at 13,000 rpm for 10 min. Plasmid DNA in the supernatant was then applied to a Qiaprep spin column from which it was purified by a series of washes and centrifugation as instructed. Finally, the DNA was eluted using 50 µl of sterile distilled water and stored at -20°C.

2.7 DNA quantification

The quantity and purity of extracted DNA was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington USA). The purity of the DNA sample was determined from the ratio of sample absorbance at 260 nm and 280 nm with ratios between 1.8 and 2.0 indicating pure DNA samples.

2.8 Primer Design and Synthesis

Primers were designed using Primer 3 input (http://frodo.wi.mit.edu/) with the aid of DNA sequences obtained via the NCBI database and their specificity ascertained using the NCBI BLAST tool.
Unless otherwise stated, primer synthesis was carried out by Sigma-Aldrich Ltd (Dorset, UK) on a 0.05 μM synthesis scale with purification by desalting. The exception to this was in the case of primers used in the HRM assays where higher purity primers are required. These were purified by HPLC.

2.9 **Polymerase Chain Reaction (PCR)**

A 10 mM dNTP solution was prepared from 100 mM stock solutions of dATP, dCTP, dGTP and dTTP (Promega, Southampton UK) by combining 50 μl of each dNTP stock with 300 μl of sterile distilled water.

Unless indicated, all DNA amplification reactions were carried out on a Techne TC-312 cycler (Barloworld Scientific, Staffordshire UK) in a 50 μl volume made up of 1 × Phusion High-Fidelity DNA Polymerase buffer (New England BioLabs, Herts, UK), 200 μM dNTPs, 0.5 μM of each primer, 0.5 μl of template DNA and 0.5 μl of 2 U/μl Phusion High-Fidelity DNA Polymerase (New England BioLabs, Herts, UK) and sterile distilled water. Thermal cycling conditions consisted of an initial denaturation at 98°C for 5 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 55°C for 30 s and elongation at 72°C for 30 s with a single final extension at 72°C for 5 min.

2.10 **DNA Electrophoresis**

DNA fragments were visualised following agarose gel electrophoresis using 1 – 1.5% w/v agarose gels prepared in 1 × TAE buffer (Appendix 2) containing 5 μg/ml.
ethidium bromide (EtBr). Aliquots of DNA samples and a 1 kb standard ladder (Promega, Southampton, UK) were added to 5 × loading buffer and then loaded into wells in the set agarose gel. Following electrophoresis for 45 – 60 min at 80 – 100 V (BioRad PowerPac 200/2.0), visualisation of dsDNA was then carried out using a UV transilluminator and the image was captured using the attached Uvitec gel image capture equipment (Uvi Tec, Cambridge, UK).

2.10.1 Purification of PCR products
Where necessary, purification of PCR products was carried out to remove excess primers and PCR components in order to prevent potential interference with downstream applications. This was done using either a QIAquick PCR Purification Kit or a QIAquick Gel Extraction Kit (Qiagen, Crawley UK) according to the manufacturer’s protocols. Elution of pure DNA was then carried out using 50 μl sterile water after a series of washes to remove contaminants.

2.10.2 DNA Sequencing
All DNA sequencing in the present study was carried out at the Biopolymer Synthesis and Analysis Unit (BSAU) of the School of Biomedical Sciences (University Of Nottingham) with 3130xl ABI PRISM Genetic Analyzer and the BigDye version 3.1, dye terminator chemistry. Sequences were analysed using Chromas Lite 2.0 software (http://chromas-lite.software.informer.com/2.0/). The primers employed for each sequencing reaction are described in the relevant result chapter of this thesis.
2.11 Methods used in DNA manipulation

2.11.1 A-tailing
In order to insert blunt ended DNA fragments generated using Phusion High-Fidelity DNA Polymerase into the pGEM®-T plasmid vector (Promega, Southampton UK), it was first necessary to create a dATP tail at both 5’ ends of the fragment. The A-tailing procedure was carried out as described in the Promega pGEM-T and pGEM-T Easy vector systems technical manual. This involved the incubation of 10 µl of purified DNA fragment in a 20 µl reaction mixture composed of 2 mM MgCl₂, 0.2 mM dATP, 2 µl Taq DNA polymerase (5 U/ml) and 1 × Taq DNA buffer at 70°C for 25 min.

2.11.2 DNA Ligation
Ligation of DNA fragments was carried out using Promega T4 DNA ligase in a 10 µl ligation mixture. This consisted of 1 × ligation buffer (30 mM Tris-HCl pH 7.8, 10 mM MgCl₂, 10 mM DTT and 1 mM ATP), 1 µl of 1000 U/ml T4 DNA ligase and varying volumes of DNA insert and vector and sterile distilled water. The ligation reaction was carried out overnight at 16°C.

Following this, the ligation mixture was dialysed for 15 min against sterile distilled water using a 0.025 µm filter (Millipore) prior to transformation or storage at -20°C.

2.11.3 Restriction Enzyme (RE) Digestion
Digestion of plasmid DNAs using XhoI and NcoI (New England Biolabs, Herts, UK) restriction enzymes was carried out in a 20 µl reaction volume containing 1 × NEB Buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate,
1 mM DTT pH7.9), 2.5 µl of DNA, 1 µl of each enzyme (20,000 U/ml) and distilled water which was incubated for 1 h at 37°C.

2.11.4 Preparation of Competent *E. coli*

In order to prepare competent cells, an overnight bacterial culture of the appropriate strain of *E. coli* grown in LB broth was diluted 1:100 into 500 ml fresh LB broth. This was incubated at 37°C with shaking until an OD$_{600nm}$ of 0.5 was reached. Bacterial cells were then centrifuged at 4°C for 15 min at 4000 × g using an Allegra X-22R (Beckman Coulter, High Wycombe, UK). Harvested cells were subsequently washed by a series of resuspensions and centrifugation steps using 1 × and 0.5 × original culture volumes of sterile cold water. Following a final wash and centrifugation step using 20 ml of cold 10% v/v glycerol in water, competent cells were recovered into 1 ml of 10% v/v glycerol in water, aliquoted, fast frozen using liquid nitrogen and stored at -80°C.

2.11.5 Transformation of *E. coli* by electroporation

10 µl of plasmid DNA was added to 40 µl of thawed electrocompetent cells on ice without mixing. This mixture was then transferred into a 2 mm cuvette (Geneflow Staffordshire, UK) and electroporated using a MicroPulser Electroporator (BioRad, Hemel Hempstead UK) by applying 2.5 kV at 200 Ω and 25 µFD. This was followed by the immediate addition of 800 µl of LB broth. Following a 2 h incubation at 37°C, 200 µl of culture was then plated onto LB agar plates supplemented with 100 µg/ml ampicillin. Plates were incubated overnight at 37°C to select for transformants.
2.12 Protein Separation and Visualisation

2.12.1 Preparation of *E. coli* cell lysates by sonication

Lysis of *E. coli* BL21/DE3 cells expressing recombinant proteins was carried out by sonication. Cell pellets resuspended in PBS (or appropriate buffer for column purification) were cooled on ice for 10 min prior to sonication. Cells were sonicated at an amplitude of 5 microns for 60 s using a Soniprep 150 (Sanyo). Sonication was repeated 3 - 5 times with cooling on ice between cycles to prevent overheating of the samples.

2.12.2 Cell Fractionation

The presence of expressed proteins in soluble and insoluble *E. coli* fractions was investigated following fractionation of sonicated whole cell lysates. Following centrifugation of lysates at 13,000 rpm for 10 min, the resulting supernatant corresponding to the soluble fraction was removed and retained. The pellet which represents the insoluble fraction was then washed in PBS to remove any remaining soluble fraction and re-centrifuged. Cell fractions were stored at -20°C.

2.12.3 Sample Preparation for SDS-PAGE

Protein samples for SDS-PAGE were denatured by a combined effect of sample buffer and heat. This was done by adding the appropriate volumes of single or double strength SDS sample buffer (Appendix 2). 2 × SDS sample buffer was added to the whole cell lysate and soluble fractions, while the insoluble fraction was resuspended in 1 × SDS sample buffer. All samples were boiled for 5 min. Samples were subsequently centrifuged at 13,000 rpm for 3 min to remove debris.
2.12.4 SDS-PAGE

SDS-PAGE is a technique which enables the separation of proteins based entirely on differences in molecular weight/mass and was initially described in 1970 (Laemmli, 1970). In the present study 15 µl of protein sample prepared as described in Section 2.12.3 was loaded into each well of a discontinuous SDS-PAGE gel (11.5% w/v acrylamide resolving gel and 4% w/v acrylamide stacking gel) using a BioRad Mini Protean II gel system and electrophoresed at 200 V for 42 min (BioRad PowerPac 200/2.0). The composition of the gels and running buffer are shown in Appendix 2. Low range molecular weight markers (BioRad, Hemel Hempstead UK) were electrophoresed on each gel.

Protein bands were subsequently visualised following fixation, by staining with Coomassie brilliant blue (BioRad, Hemel Hempstead UK) according to the manufacturer’s instructions and destaining with Methanol/Acetic acid (Appendix 2).
Chapter Three

3 Molecular characterization and PVL typing of community-acquired meticillin-sensitive *S. aureus* clinical isolates

3.1 Introduction

Following the firmly established epidemiological link between PVL and CA-MRSA (Chheng et al., 2009, Diep et al., 2004, Enany et al., 2007, Liassine et al., 2004, Linde et al., 2005, Miklasevics et al., 2004, Tsai et al., 2008, Vandenesch et al., 2003, Witte et al., 2005), and the subsequent reports of worldwide spread, significant research has been aimed at the typing and characterisation of PVL-positive *S. aureus* isolates. Despite the fact that PVL-MSSA has been reported to pose an equal clinical threat as PVL-MRSA (Schefold et al., 2007, Mongkolrattanothai et al., 2003, Perbet et al., 2010, Swaminathan et al., 2006, Lorenz et al., 2007, Ramos et al., 2009), most research has however focused on PVL-MRSA rather than PVL-MSSA. A general lack of knowledge therefore exists about this subset of isolates, especially with regards to its genotypic characteristics. In the past 5 years, the rise in number of PVL-positive *S. aureus* in the United Kingdom (as referred to the HPA) has surprisingly been due to meticillin susceptible rather than resistant isolates (HPA, 2011a). This observation has also been reported in several other countries (Tong et al., 2010, Breurec et al., 2011). PVL-MSSA may therefore present a previously unrecognised public health risk. In addition to this, a second possible role exists for this group of isolates in respect of evolution of *S. aureus*. 
CA-MRSA are generally accepted to have evolved following the acquisition of SCCmec elements by strains of *S. aureus* in the community (Daum et al., 2002). For the actual evolution of PVL-MRSA however, two possibilities exist (Figure 3.1). Uptake of genes encoding PVL by these community strains could either have occurred prior to the acquisition of the SCCmec elements or these genes might have been introduced after SCCmec acquisition. The acquisition of PVL encoding genes prior to SCCmec acquisition is however the more widely accepted model for the evolution of PVL-MRSA (Monecke et al., 2007b). This view has often been supported by studies describing PVL-MSSA strains which are highly similar to PVL-MRSA strains but differ in SCCmec carriage (Boyle-Vavra and Daum, 2007, Mongkolrattanothai et al., 2003, Monecke et al., 2007b, Diep et al., 2006a). In addition, a significantly higher genetic diversity has been reported within PVL-MSSA than in PVL-MRSA (Vandenesch et al., 2003, Rasigade et al., 2010). This finding further supports the possible emergence of PVL-MRSA via PVL-MSSA as opposed to the PVL-MSSA strains resulting from a possible loss of a SCCmec element. However, the hypothesis that PVL-MSSA isolates serve as precursors for PVL-MRSA has not yet been proven.
Figure 3.1: Hypothetical pathways for the evolution of PVL-MRSA

Acquisition of genes encoding PVL by the MSSA strain shown on the left of the figure could either occur 1. Prior to SCCmec acquisition as shown in the top half of the figure or 2. Subsequent to SCCmec acquisition, as shown in the bottom half of the figure.
Very few studies have however focused on the genetic background of PVL-MSSA strains, PVL phage variant carried, and variations within the *lukSF-PV* locus, to see how these all fit in with these evolutionary theories. Some studies simply characterised the isolates using MLST and/or *spa* typing (Udo et al., 2008, Holmes et al., 2005, Nimmo et al., 2006) and others went on to look at the virulence gene profiles and antibiogram (Tristan et al., 2007, Witte et al., 2005, Bartels et al., 2007, Denis et al., 2005). However, PVL gene polymorphisms and their possible implications had not been addressed at the start of the present study. Being able to show that a PVL-MSSA strain shares not just the same MLST and *spa* profile as its MRSA counterpart, but also contains identical PVL sequence variation and phage types would prove invaluable to understanding the evolution of these organisms. This represents a key aim of the work described in this chapter.

In addition to contributing to the limited available molecular typing data on PVL-MSSA strains, the work described in this chapter set out to explore PVL gene polymorphisms and phage distribution in a group of clinical isolates and how these relate to those previously observed in PVL-MRSA strains. This would aid current understanding of the evolution and emergence of PVL-positive CA-MRSA isolates and help to more accurately assess the current threat posed by these often overlooked PVL-positive MSSA strains.
3.2 Materials and Methods

3.2.1 Bacterial Isolates

Test isolates used in the present study were previously characterized clinical isolates of *S. aureus* collected by Dr. Tim Sloan from Nottingham University Hospitals NHS Trust (NUH) (Table 2.1). The clinical isolates were selected as likely to be PVL producers based upon (a) clinical suspicion due to the type of infection, or (b) an antibiotic susceptibility profile (gentamicin/trimethoprim resistance) that had previously been found to be characteristic of PVL-MSSA strains in NUH. The presence or absence of thirteen toxin genes (*sea, seb, sec, sed, see, seg, seh, sei, sej, tst, eta, etb* and *etd*) in these isolates had previously been determined by the Health Protection Agency (HPA) Staphylococcus Reference Unit (SRU), using multiplex PCRs as previously described (Becker et al., 1998, Monday and Bohach, 1999).

3.2.2 Strain Characterisation

To determine the relatedness and genetic origins of these isolates, sequence based molecular techniques were employed to determine the *spa* type, MLST type and presence of allelic variations in the two genes that encode the PVL toxin. In addition, characterisation of the PVL phage type was carried out. All of these methods are similar in comprising a PCR DNA amplification using specific primers, amplicon purification, quantification and a sequencing step.
Genomic DNA employed in these tests was extracted from overnight cultures of test isolates grown in LB broth using the Qiagen DNeasy® blood and tissue kit as previously described (Section 2.5).

3.2.2.1 spa typing

The first step in spa typing is the amplification of the polymorphic X region of the spa gene. This was carried out as previously described (Section 2.9) using primers described in a previous study (Aires-de-Sousa et al., 2006) and listed in Table 3.1. Following purification and quantification, the PCR products were sequenced using the same forward and reverse primers as used for amplification. Analysis of the spa gene sequence was performed manually, based on the process described by Harmsen and colleagues (Harmsen et al., 2003), using data presented on the Ridom GmbH spa website (www.spaserver.ridom.de).

The first step in this process involved the identification of specific 5′ and 3′ signature sequences (http://spa.ridom.de/submission.shtml). These signature sequences serve as markers for the start and end of the polymorphic X region used in spa typing and ensure that no repeat is omitted. Following the identification of the signature sequences, a manual analysis of the test sequence was carried out in order to identify the individual repeats. First, the 24 bases immediately adjacent to the 5′ signature sequence were cross-matched against the spa repeats database (http://spa.ridom.de/repeats.shtml) to assign the repeat identities. This process was repeated until the ‘ending’ repeat was identified. The ‘ending’ repeat is the last repeat and this can be found 18 – 19 bases from the 3′ signature sequence. Though not a
common occurrence, repeat sequences 21 bases and 27 bases in length have also been reported (http://spa.ridom.de/repeats.shtml). This possibility was therefore taken into consideration during the cross-matching process.

Next, spa types were assigned for each test sequence. A spa type is a unique identifier given to a particular repeat succession. For example, spa type 001 (t001), is the unique identifier for the repeat succession 26-30-17-34-17-20-17-12-17-16. In the present study, spa types were assigned by manually cross-matching the repeat succession for each test sequence against the spa-type database (http://spa.ridom.de/spatypes.shtml). In the case of a suspected novel spa type, i.e. where the spa sequences failed to match up with any of the previously described spa types, the sequence chromatogram was then submitted to the spaserver (http://spa.ridom.de/submission.shtml). This was to enable the spa curator to verify the accuracy of the reads, presence of the signature sequences and to assign a novel spa type identity directly from the chromatogram.
### Table 3.1: Primers Used In DNA amplification in the present study

<table>
<thead>
<tr>
<th>Target</th>
<th>5’ – 3’ Primer sequences</th>
<th>Product Length (kb)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>spa</td>
<td>Forward: TAAAGACGATCCTCTCGGTGAGC Reverse: CAGCAGTAGTGCCGGTTTGT</td>
<td>Variable</td>
<td>(Aires-de-Sousa et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLST</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arc</td>
<td>Forward: TTGATTCACCAGCAGCGTTGATGTC Reverse: AGGTATCTGCTCTTCAATCAGCG</td>
<td>570</td>
<td>(Enright et al., 2000)</td>
</tr>
<tr>
<td>aroE</td>
<td>Forward: ATCGGAATCCTATTTCAATTC Reverse: GGTGTTGATATTAAATAACGATATC</td>
<td>536</td>
<td></td>
</tr>
<tr>
<td>glpF</td>
<td>Forward: CTAGGAACCTGCAACTTTAATCC Reverse: TGGTAAAATCGCATGTCCAATTC</td>
<td>576</td>
<td></td>
</tr>
<tr>
<td>gmk</td>
<td>Forward: ATCGTTTTTATCGGGACCACC Reverse: TCAATATTACAAACGTAACGTA</td>
<td>488</td>
<td></td>
</tr>
<tr>
<td>pta</td>
<td>Forward: GTTAAAATCGTATTACCTGAAGG Reverse: GACCCCTTTTGTGAAAAGCTTAA</td>
<td>575</td>
<td></td>
</tr>
<tr>
<td>tpi</td>
<td>Forward: TCGTTCATTCTGAACGTGCTGAA</td>
<td>475</td>
<td></td>
</tr>
<tr>
<td>yqiL</td>
<td>Forward: CAGCATACAGGACACCTATTTGGA Reverse: CGTTGAGGAATCGATCTGGAAAC</td>
<td>598</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lukS</td>
<td>Forward: ATGGTCAAAAAAAGACTATTAGCTG Reverse: TCAAATCTACTGTTTCTCTTGAG</td>
<td>764</td>
<td>The present study</td>
</tr>
<tr>
<td>lukF</td>
<td>Forward: TCAATACACTGTTGATCTCTCCTGAG</td>
<td>535</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lukS</td>
<td>Forward: GTGGTCCATCAAACAGGAGGT Reverse: TGAAGGATTGGAACACTGTTTG</td>
<td>272</td>
<td>The present study</td>
</tr>
<tr>
<td>lukF</td>
<td>Forward: CGGTAGGTTATCTTATGGTGGA Reverse: TCCAGTGATGTTGATCCAAAA</td>
<td>447</td>
<td>The present study</td>
</tr>
</tbody>
</table>
3.2.2.2 Multilocus Sequence Typing (MLST)

Selected isolates were further typed using the MLST technique which involves the amplification and sequence comparison of fragments of seven housekeeping genes as previously described (Enright et al., 2000), using primers listed in Table 3.1. In brief, DNA amplification was carried out at each locus. PCR products were then purified, quantified and sequenced as previously described. Each sequence was edited to correspond to the specific region of internal fragment used to define an allele using information contained in the *S. aureus* MLST database (http://saureus.mlst.net/sql/concatenate/default.asp). Next, the allelic profile for each isolate was assigned by comparing the sequence data for each locus against sequences contained in the *S. aureus* MLST database (http://saureus.mlst.net/sql/multiplelocus.asp). Finally the unique sequence type represented by each allelic profile was determined by comparing against the allelic profiles contained in the *S. aureus* MLST database (http://saureus.mlst.net/sql/allelicprofile_choice.asp).

Alternatively, sequence types were established from *spa* types via the *spa* database http://spa.ridom.de/mlst.shtml, which enables mapping of MLSTs based on *spa* types.

3.2.2.3 PVL characterization and PVL phage typing

As PVL typing simply involves an analysis of specific allelic variations on both the *lukF* and *lukS* genes, two sets of primers (Table 3.1) were designed to amplify the
variable regions in the *lukSF-PV* locus. The first set of primers was designed to generate a 764 bp fragment of the *lukS* gene starting at nucleotide 1 of the *lukSF-PV* locus. The second set of primers was designed to generate a 535 bp fragment of the *lukF* gene starting at nucleotide 1,288 of the *lukSF-PV* locus. Amplification products were generated using the same cycling conditions as previously described (Section 2.9) and purified PCR products sequenced. The PVL types were determined based on the gene sequence (O’Hara et al., 2008).

PVL phage typing was carried out as previously described (Ma et al., 2008), using additional information provided by Boakes and colleagues (Boakes et al., 2011a). In brief, this method designed to detect the six PVL encoding phages so far described (øPVL, ø108PVL, ø2958PVL, øSLT, øSa2mw and øSa2USA), involves the use of nine different amplification reactions (PCR-1 to PCR-9). Primers for PCR-1 to PCR-8 were described by Ma and colleagues. These were based on the nucleotide sequences of five PVL phages (øPVL, ø108PVL, ø2958PVL, øSLT and øSa2mw) with GenBank accession numbers of AB009866, AB009866, AP009363, NC_002661 and BA000033, respectively.

PCR-1 and PCR-2 were specifically designed to differentiate the different phage morphologies. PCR-1 (Table 3.2) was composed of two sets of primers specific for DNA fragments unique to the icosahedral-head type PVL phages. The two sets of primers in PCR-2 were however specific for DNA fragments unique to PVL phages with elongated-head types. PCR-3 and PCR-4 were then designed to amplify specific linkage genes to the *lukSF-PV* locus. These reactions were to confirm that the phage
morphologies detected were indeed associated with carriage of the genes encoding for PVL. PCR-3 detected linkage genes in icosahedral-head type PVL phages while PCR-4 detected linkage genes in elongated-head type PVL phages. Finally, the last five amplification reactions (PCR-5 to PCR-9), were designed to detect phage specific fragments. PCR-6, PCR-7, PCR-8 and PCR-9 detected DNA fragments unique to ø2958PVL, øSLT, øSa2mw and øSa2USA respectively. PCR-5 however, was designed to identify both øPVL and ø108PVL. While this reaction contained primers which could amplify phage specific fragments from both of these phages, the product lengths were however distinct for each phage (Figure 3.2).

These nine amplification reactions were expected to generate amplicons ranging in length from 468 bp to 10,497 bp. The amplification protocol used was similar to that previously described (Section 2.9) but had a longer elongation step (2 min), thirty five cycles and DMSO was added to reaction mixtures to facilitate generation of the longer product lengths expected.
Table 3.2: Primers used for detecting and typing PVL phages

<table>
<thead>
<tr>
<th>Reaction ID</th>
<th>Primer name</th>
<th>5′ – 3′ Primer Sequence</th>
<th>Product Length (bp)</th>
<th>Phages Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR-1</td>
<td>portal-1F</td>
<td>ACACGTGATAAAAACAGGAGAA</td>
<td>569</td>
<td>ø108PVL øPVL</td>
</tr>
<tr>
<td></td>
<td>portal-1R</td>
<td>TCTAAATTAGCATCCGTGATAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tail-1F</td>
<td>ATAAATGGGATAGCAACGCAA</td>
<td>489</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tail-1R</td>
<td>CTTGATTAGACTCAACCAAACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-2</td>
<td>portal-2F</td>
<td>GATGGCTAGTGGCCTTGA</td>
<td>656</td>
<td>ø2958PVL øSa2mw øSLT</td>
</tr>
<tr>
<td></td>
<td>portal-2R</td>
<td>CTGGAGGCAATGGAAAAACG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tail-2F</td>
<td>CATAGCGCTAATGTCCGAA</td>
<td>468</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tail-2R</td>
<td>AGCCTCCATTGTTGTTTGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-3</td>
<td>lukSR1</td>
<td>ACGAAGTAGCAATAGGAGTGA</td>
<td>10,497</td>
<td>ø108PVL øPVL</td>
</tr>
<tr>
<td></td>
<td>teil-ico-F</td>
<td>AGATTAGAAGAGGAGGACGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-4</td>
<td>lukSR1</td>
<td>ACGAAGTAGCAATAGGAGTGA</td>
<td>9,483</td>
<td>ø2958PVL øSa2mw øSLT</td>
</tr>
<tr>
<td></td>
<td>teilE-F2</td>
<td>ATGATTCAAACTGTITTCTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-5</td>
<td>intF2</td>
<td>ATGTTTTCCGAGTTTTTGAGTTAG</td>
<td>4,340</td>
<td>ø108PVL øPVL</td>
</tr>
<tr>
<td></td>
<td>108-aR</td>
<td>TCAAATCCGTAATCCTCATTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVL-aR</td>
<td>TTCACTAACTAACCCTATCTTCT</td>
<td>1,411</td>
<td></td>
</tr>
<tr>
<td>PCR-6</td>
<td>intF2</td>
<td>ATGTTTTCCGAGTTTTTGAGTTAG</td>
<td>2,238</td>
<td>ø2958PVL</td>
</tr>
<tr>
<td></td>
<td>2958-aR</td>
<td>TGGTAATCAACCATTCTTATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-7</td>
<td>intF2</td>
<td>ATGTTTTCCGAGTTTTTGAGTTAG</td>
<td>4,065</td>
<td>øSa2mw</td>
</tr>
<tr>
<td></td>
<td>MW2-aR</td>
<td>TAAGTTTCTGGTGTCATTCTAAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-8</td>
<td>intF2</td>
<td>ATGTTTTCCGAGTTTTTGAGTTAG</td>
<td>8,770</td>
<td>øSLT</td>
</tr>
<tr>
<td></td>
<td>SLT-aR</td>
<td>TCTTACCAAATGCAACACCAAACA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR-9</td>
<td>Sa2USA_F</td>
<td>GGTATTACCCAACAAACCAAAATTACG</td>
<td>680</td>
<td>øSa2USA</td>
</tr>
<tr>
<td></td>
<td>Sa2USA_R</td>
<td>CCTCAGGGCGCCATCACCAATA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The primer sequences were compiled from (Ma et al., 2008) and (Boakes et al., 2011a)
Figure 3.2: A representative example of expected product sizes for PVL phage amplification reactions.

The primers used in each reaction are as described in Table 3.2. The PCR products were generated as described in Section 3.2.2.3. The expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5\(^a\): 4,340; PCR-5\(^b\): 1,411; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770.

PCR-5 is designed to amplify phage specific regions of two different phages. \(^a\)4,340 bp product expected when reaction is positive for \(\varnothing 108\)PVL. \(^b\)1,411 bp product expected when reaction is positive for \(\varnothing PVL\).

The figure was reproduced from Ma et al., (2008).
Following amplification of the phage genes, determination of the specific phage present could then be made based on the unique combination of amplicons obtained (Table 3.3). Phages which could not be completely typed could however still be classified as either icosahedral or elongated head types.

**Table 3.3: Key to PVL phage typing**

<table>
<thead>
<tr>
<th>S/No</th>
<th>Phage</th>
<th>PCR 1</th>
<th>PCR 2</th>
<th>PCR 3</th>
<th>PCR 4</th>
<th>PCR 5</th>
<th>PCR 6</th>
<th>PCR 7</th>
<th>PCR 8</th>
<th>PCR 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>øPVL</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2.</td>
<td>ø108PVL</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3.</td>
<td>øSa2958</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4.</td>
<td>øSa2mw</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5.</td>
<td>øSLT</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>øSa2USA</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

The table shows the various PVL phages and the unique combination of positive results which indicate each phage

+: Represents the presence of amplification products of the expected sizes  
+*: Represents \( \approx 1.4 \) kb PCR-5 amplification product  
+**: Represents \( \approx 4.3 \) kb PCR-5 amplification product  
-: Represents a lack of amplification product of expected size

Expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5: 4,340 or 1,411; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770.
3.3 Results

3.3.1 Characteristics of PVL-MSSA

The 19 PVL-positive MSSA strains analysed comprised both hospital and community-acquired isolates recovered between November 2008 and May 2009 at NUH. Isolates were obtained from a range of sample types, from patients with varying clinical histories and ages (Table 3.4). These isolates were mostly community-acquired (15/19; 80%), i.e. either sent from the GP or recovered at A/E and were associated mainly with skin and soft tissue infection. One strain (TS1) however was isolated from a fatal case of necrotizing pneumonia. This strain collection was noted to have a low prevalence of virulence/toxin genes. While 89.5% (17/19) of isolates carried the seg and sei genes, 31.6% (6/19) of isolates also carried one or two additional toxin genes (sea, sec or sed). Two isolates carried four of the thirteen toxin genes investigated (Table 3.4). The majority of toxin genes (seb, see, seh, sej, tst, eta, etb and etd) were however not detected in this population.

89.5% (17/19) of isolates were trimethoprim resistant with additional gentamicin resistance observed in 57.9% (11/19) of isolates. This high prevalence of gentamicin/trimethoprim resistance was probably a reflection of the selection criteria. Most isolates (11/19; 57.9%) were resistant to a maximum of 2 antibiotics. The highest resistance (to four antibiotics) was observed in a single ST30 isolate (TS12). All the isolates were susceptible to clindamycin, rifampicin, linezolid, vancomycin, fusidic acid and teicoplanin.
Table 3.4: Clinical characteristics of Panton-Valentine leucocidin positive *S. aureus* test isolates.

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID /Source of Isolation</th>
<th>Toxin gene profile</th>
<th>Age</th>
<th>Clinical details</th>
<th>Antibiotic resistances c, d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS 1/CA</td>
<td>None</td>
<td>30</td>
<td>Necrotising pneumonia, fatal</td>
<td>FULLY SENSITIVE e</td>
</tr>
<tr>
<td>2.</td>
<td>TS 5/CA</td>
<td><em>sea, sei, sec, seg</em></td>
<td>39</td>
<td>Axillary boil</td>
<td>TMP, CIP, ERY</td>
</tr>
<tr>
<td>3.</td>
<td>TS 6/HA</td>
<td><em>sei, seg</em></td>
<td>80</td>
<td>Wound infection post hip surgery</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>4.</td>
<td>TS 7/CA</td>
<td><em>sea, sei, seg</em></td>
<td>35</td>
<td>Nasal carriage</td>
<td>TMP, DOX</td>
</tr>
<tr>
<td>5.</td>
<td>TS 8/CA</td>
<td><em>sea, sei, sec, seg</em></td>
<td>17</td>
<td>Facial abscess</td>
<td>TMP, ERY, CIP</td>
</tr>
<tr>
<td>6.</td>
<td>TS 9/HA</td>
<td><em>sei, seg</em></td>
<td>89</td>
<td>Wound infection post hip surgery</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>7.</td>
<td>TS 12/CA</td>
<td><em>sei, seg</em></td>
<td>6</td>
<td>Wound, no details given</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>8.</td>
<td>TS 13/CA</td>
<td><em>sei, seg</em></td>
<td>35</td>
<td>Leg abscess</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>9.</td>
<td>TS 14/HA</td>
<td><em>sei, seg</em></td>
<td>87</td>
<td>Leg ulcer</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>10.</td>
<td>TS 15/CA</td>
<td><em>sei, seg</em></td>
<td>39</td>
<td>Recurrent boils</td>
<td>FULLY SENSITIVE</td>
</tr>
<tr>
<td>11.</td>
<td>TS 16/CA</td>
<td><em>sea, sei, seg</em></td>
<td>27</td>
<td>Thigh abscess</td>
<td>TMP</td>
</tr>
<tr>
<td>12.</td>
<td>TS 17/CA</td>
<td><em>sei, seg</em></td>
<td>55</td>
<td>Nasal carriage</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>13.</td>
<td>TS 18/CA</td>
<td><em>sei, seg</em></td>
<td>83</td>
<td>Foot wound</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>14.</td>
<td>TS 19/CA</td>
<td><em>sei, seg</em></td>
<td>26</td>
<td>Back abscess</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>15.</td>
<td>TS 20/CA</td>
<td><em>sei, seg</em></td>
<td>64</td>
<td>Bursitis, knee</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>16.</td>
<td>TS 21/HA</td>
<td>None</td>
<td>3</td>
<td>Empyema</td>
<td>TMP</td>
</tr>
<tr>
<td>17.</td>
<td>TS 23/CA</td>
<td><em>sei, seg</em></td>
<td>24</td>
<td>IVDU, cellulitis, bacteraemia</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>18.</td>
<td>TS 24/CA</td>
<td><em>sei, seg</em></td>
<td>83</td>
<td>Cellulitis foot, bacteraemia</td>
<td>GEN, TMP</td>
</tr>
<tr>
<td>19.</td>
<td>TS 25/CA</td>
<td><em>sei, sed, seg</em></td>
<td>55</td>
<td>Severe pneumonia, post-viral</td>
<td>TMP</td>
</tr>
</tbody>
</table>

aSource of Isolation, HA, isolated after 48 h of admission; CA, isolated either at the general practice or accident and emergency department or within 48 h of admission

bTotal toxin genes analysed for include: *sea, seb, sec, sed, see, seg, seh, sei, sej, tst, eta, etb* and *etd*

dAll isolates were susceptible to clindamycin, rifampicin, linezolid, vancomycin, fusidic acid and teicoplanin
eFully sensitive to the listed antibiotics
3.3.2 *spa* typing of test isolates

Following amplification of the *spa* gene from various isolates, and purification using a QiaQuick PCR purification kit, DNA sequencing resulted in noisy chromatograms (Figure 3.3) as defined by irregularly spaced reads. Multiple peaks were observed for a single base in some cases. This was considered unacceptable as sequences could not be accurately assigned. To overcome this problem, amplicons were instead purified following electrophoresis by gel extraction using a QiaQuick gel extraction kit (Section 2.10.1) prior to sequencing. This eliminated the DNA smearing which had been previously observed when products were analysed (Figure 3.4) and resulted in clean sequencing reads with reduced baseline noise (Figure 3.5).
Figure 3.3: Example of a noisy sequence chromatogram generated following sequencing of the spa amplicons purified with a QiaQuick PCR purification kit.

The sequence data was generated (using the recommended spa typing primers listed in Table 3.2) following purification of a spa amplicon using the QiaQuick PCR purification kit as described in Section 2.10.1. The arrows indicate examples of multiple peaks generated for a single nucleotide resulting in a noisy chromatogram producing a potentially ambiguous sequencing result.
Figure 3.4: Representative gel images indicating purity of the spa gene fragments following amplicon purification by gel extraction.

The agarose gel images were generated following DNA electrophoresis of purified spa gene fragments as described in Section 2.10. The left hand gel image (Panel A), shows a smearing (as indicated by the arrows) associated with several spa gene fragments following purification using a QiaQuick purification kit i.e. directly on amplified DNA as described in Section 2.10.1. The right hand image (Panel B), shows the lack of smearing associated with results of purification of spa gene fragments using a QiaQuick extraction kit as described in Section 2.10.1.
Figure 3.5: Example of a clean sequence chromatogram generated following sequencing of the *spa* amplicons purified with a QiaQuick gel extraction kit.

The sequence data was generated (using the recommended *spa* typing primers listed in Table 3.2) following purification of a *spa* amplicon using the QiaQuick gel extraction kit as described in Section 2.10.1. A clean sequence chromatogram is indicated by the lack of multiple peaks and reduced baseline to noise.
For each *spa* sequence, repeat identities and *spa* types were mapped only after the detection of the 5’ and 3’ signature sequences that are essential for reliable results (http://spa.ridom.de/submission.shtml). An example of the elements identified following sequence analysis is shown in Figure 3.6. The results of the *spa* typing showed that eleven different *spa* types were present in the strain collection. The number of repeat sequences ranged from five to twelve (Table 3.5). These eleven types clustered into two *spaCCs* (CC005 and CC345/657) and five singletons based on the BURP algorithm (StaphType Software; Ridom GmbH, Wurzburg, Germany).

5' signature – 0 bp from repeat

3' signature – 19bp from repeat

Figure 3.6: Representation of sequence analysis and the spa type allocation process for TS13

The spa sequence was generated following amplification, purification and sequencing of a spa gene fragment (Figure 3.2.2.1). The assignment of spa types depends on the presence of signature sequences (highlighted in yellow) which serve as markers for the start and end of the polymorphic X region used in spa typing. The spa repeats identified by cross-matching each 24 base cluster against the spa repeats database (http://spa.ridom.de/repeats.shtml) are represented as figures in brackets.
Types t005 and t021 occurred most frequently (42.1%) while seven other spa types were represented by only a single clinical isolate. Three spa types which have not previously been described were identified and submitted to the Ridom SpaServer (http://spa.ridom.de/submission.shtml), and assigned as spa types t6642, t6643 and t6769. Both t6642 and t6643 were found to be associated with ST22 isolates. t6642 with a repeat succession of 26-23-13-23-31-05-22-17-25-17-25-114 was of the same length as the prototype t005 (identified by the BURP algorithm). These however varied in the composition and arrangement of repeat sequences (t005 had a repeat succession of 26-23-13-23-31-05-17-25-17-25-16-28). Type t6643 (26-23-13-23-31-05-17-25-16-390) however differed in length with 2 repeat deletions noted. Similarly, type t6769 was of the same length as type t186 (an ST88 related spa type) with single point mutations occurring in three repeats leading to an r34 to r13 change (Figure 3.7).
Table 3.5: Results of *spa* typing.

<table>
<thead>
<tr>
<th>S/No</th>
<th>Test Isolate</th>
<th>Repeat Succession&lt;sup&gt;a&lt;/sup&gt;</th>
<th><em>spa</em> Type&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. TS13, TS14, TS18, TS20, TS24</td>
<td>26-23-13-23-31-05-17-25-17-25-16-28</td>
<td>t005</td>
<td></td>
</tr>
<tr>
<td>2. TS7, TS12, TS15, TS16</td>
<td>15-12-16-02-16-02-25-17-24</td>
<td>t021</td>
<td></td>
</tr>
<tr>
<td>3. TS19</td>
<td>26-23-31-05-17-25-17-25-16-28</td>
<td>t310</td>
<td></td>
</tr>
<tr>
<td>4. TS8</td>
<td>26-23-13-21-17-34-33-34</td>
<td>t345</td>
<td></td>
</tr>
<tr>
<td>5. TS2, TS5</td>
<td>26-23-13-21-17-34-33-34</td>
<td>t657</td>
<td></td>
</tr>
<tr>
<td>6. TS17</td>
<td>07-23-13-23-31-05-17-25-17-25-16-28</td>
<td>t852</td>
<td></td>
</tr>
<tr>
<td>7. TS1</td>
<td>07-56-12-17-17-16-16-33-31-57-12</td>
<td>t1941</td>
<td></td>
</tr>
<tr>
<td>8. TS25</td>
<td>07-23-21-17-13-33-13</td>
<td>t3342</td>
<td></td>
</tr>
<tr>
<td>9. TS6, TS9</td>
<td>26-23-13-23-31-05-22-17-25-17-25-114</td>
<td>t6642</td>
<td></td>
</tr>
<tr>
<td>10. TS23</td>
<td>26-23-13-23-31-05-17-25-16-390</td>
<td>t6643</td>
<td></td>
</tr>
<tr>
<td>11. TS21</td>
<td>07-12-21-17-13-13-13-33-13</td>
<td>t6769</td>
<td></td>
</tr>
</tbody>
</table>

The *spa* typing data presented in this table was generated as described in **Section 3.2.2.1**. The table shows the *spa* repeat succession detected in each isolate and the corresponding *spa* types.

<sup>a</sup>Repeat succession represents the sequence of repeats as identified from the *spa* repeats database ([http://spa.ridom.de/repeats.shtml](http://spa.ridom.de/repeats.shtml)).

<sup>b</sup>The *spa* type represents a unique identifier for each specific repeat succession as determined from the *spa*-type database ([http://spa.ridom.de/spatypes.shtml](http://spa.ridom.de/spatypes.shtml)).
Figure 3.7: Depiction of the relationship between the new *spa* types described in the present study and their existing clonal prototypes as defined by the BURP algorithm

The relationship between the new *spa* types t6642 and t6643 and the existing clonal prototype t005, and between t6769 and the clonal prototype t186 is shown. Yellow highlights indicate repeat variations, whilst red highlights depict the single nucleotide variation between repeats r13 and r34.
3.3.3 Multilocus Sequence Typing

Typing of the clinical isolates using the MLST technique (Enright et al., 2000) revealed further strain diversity. Six multilocus STs (ST1, ST22, ST30, ST88, ST772 and ST1518) were detected (Table 3.6). These were grouped by eBURST analysis (Section 1.8.1) (Feil et al., 2004) into five CCs of known MRSA lineages (CC1, CC22, CC30, CC88 and CC152). Of these, ST1518 was described for the first time ever in the present study. The associated isolate (TS1) was isolated from a fatal case of necrotising pneumonia. ST1518 is a single locus variant of ST152, differing by a single mutation in the glp allele. ST22 isolates occurred most frequently (9/19; 47.4%) in this collection of clinical isolates.
Table 3.6: MLST allelic profiles and corresponding STs

<table>
<thead>
<tr>
<th>S/No</th>
<th>Test Isolate</th>
<th>Allelic Profile&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sequence Type&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS25</td>
<td>1-1-1-1-1-1-1-1</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>TS6, TS9, TS13, TS14, TS17, TS18, TS19, TS20, TS23, TS24</td>
<td>7-6-1-5-8-8-6</td>
<td>22</td>
</tr>
<tr>
<td>3.</td>
<td>TS7, TS12, TS15, TS16</td>
<td>2-2-2-2-6-3-2</td>
<td>30</td>
</tr>
<tr>
<td>4.</td>
<td>TS21</td>
<td>22-1-14-23-12-4-31</td>
<td>88</td>
</tr>
<tr>
<td>5.</td>
<td>TS5, TS8</td>
<td>1-1-1-1-22-1-1</td>
<td>772</td>
</tr>
<tr>
<td>6.</td>
<td>TS1</td>
<td>46-75-207-44-13-68-60</td>
<td>1518</td>
</tr>
</tbody>
</table>

MLST of clinical isolates was carried out as described in Section 3.2.2.2

<sup>a</sup> Allelic profiles were determined based on the DNA sequence of each locus using the *S. aureus* MLST database (http://saureus.mlst.net/sql/multiplelocus.asp)

<sup>b</sup> Sequence Type corresponding to each specific allelic profile was determined from the *S. aureus* MLST database (http://saureus.mlst.net/sql/allelicprofile_choice.asp)
3.3.4 SNP analysis of the PVL-encoding genes

When compared with the sequence of øSLT, the proposed lukSF-PV progenitor (Wolter et al., 2007), a total of seven SNPs were noted in the PVL encoding gene sequences generated in the present study. Five of these occurred in the lukS gene (Table 3.7). Most isolates (17/19; 89.5%) generated sequences of the H variant as previously defined (O'Hara et al., 2008), with both H1 and H2 isoforms present among this panel of isolates. The non-synonymous A – G SNP which occurs at nucleotide 527 and defines the R variant, occurred only in a single isolate (TS1). Similar to previous reports as summarised in Table 1.4, the resulting PVL SNP profiles were generally multilocus ST specific. The H2 variant of the genes encoding PVL was detected in all ST22 isolates despite their epidemiological differences. This relationship of PVL SNP profile to multilocus ST is not exclusive as demonstrated by the detection of the H2 variant in both ST772 isolates. Two different PVL SNP profiles (H1 and H2) were however detected in ST30 isolates. The outlier H1 variant carried by TS12 was identical to that of ST22 and ST88 and overall was found in 63.2% (12/19) of this panel of isolates.
Table 3.7: Single nucleotide polymorphisms detected in the *lukSF-PV* locus of the study population

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID (ST)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Isoforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS 1 (1518)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2.</td>
<td>TS 5 (772)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>3.</td>
<td>TS 6 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>4.</td>
<td>TS 7 (30)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>5.</td>
<td>TS 8 (772)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>6.</td>
<td>TS 9 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>7.</td>
<td>TS 12 (30)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>8.</td>
<td>TS 13 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>9.</td>
<td>TS 14 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>10.</td>
<td>TS 15 (30)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>11.</td>
<td>TS 16 (30)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>12.</td>
<td>TS 17 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>13.</td>
<td>TS 18 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>14.</td>
<td>TS 19 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>15.</td>
<td>TS 20 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>16.</td>
<td>TS 21 (88)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>17.</td>
<td>TS 23 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>18.</td>
<td>TS 24 (22)</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>19.</td>
<td>TS 25 (1)</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>

The SNPs in the genes encoding PVL were detected as described in Section 3.2.2.3. The SNPs highlighted in yellow indicate point mutations that were used in defining the H1, H2 and R PVL sequence isoforms.
3.3.5 PVL Phage typing

Initial PCR amplification of phage sequences resulted in inconsistent amplification of the linkage genes (PCR-3 and PCR-4) which were expected to generate long PCR products (≈10 kb). This inconsistency was however found to be rectified by the addition of 2% v/v dimethyl sulfoxide (DMSO) to the PCRs as indicated using two independent isolates (Figure 3.8). DMSO is an organic solvent which has been demonstrated to improve PCR amplification based on its ability to prevent secondary structure formation (Jensen et al., 2010). This analysis shows the absence of PCR products of approximately 10,000 bp without DMSO addition (Figure 3.8 1A and Figure 3.8 2A) and their subsequent generation on addition of DMSO (Figure 3.8 1B and Figure 3.8 2B).
Figure 3.8: Representative gel images depicting the effect of DMSO on generation of amplicons for PVL phage typing.

The figure shows PCR amplification products generated following PVL phage typing as described in Section 3.2.3.3. Panel 1A shows the amplification products generated from TS14 genomic DNA, with no DMSO added. Panel 1B shows the amplification products generated from TS14 genomic DNA with DMSO added. Panel 2A shows the amplification products generated from TS18 genomic DNA with no added DMSO and Panel 2B shows the amplification products generated from TS18 genomic DNA with DMSO added.

The arrows on panel 1B and panel 2B indicate the extra PCR amplification product generated on addition of DMSO.

Expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5: 1,411 or 4,340; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770; PCR-9: 680.

DMSO: Dimethyl sulfoxide
A common feature of the PVL phage amplification process in the present study, was the presence of products of unexpected size as exemplified by the faint, circled bands seen in Figure 3.9. A possible lack of specificity of the primers employed in the assay was considered and the primer sequences were assessed using both a primer BLAST analysis (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and in silico PCR analysis (http://insilico.ehu.es/PCR/) against the collection of staphylococcal DNA sequences held at the National Center for Biotechnology Information (NCBI). Both tools are designed to provide theoretical PCR results. Results of in silico PCR however correlate with high specificity amplification reactions while the primer BLAST may provide information on non-specific products generated due to primer mismatches (Bikandi et al., 2004, Ficetola et al., 2010). These analyses provided a possible explanation for the presence of the faint products. Using the more stringent in silico PCR analysis which uses sequenced bacterial genomes and allows for only one mismatch, a single primer pair (PCR-6) was found to generate an unexpected 1,307 bp product. This was confirmed following a primer BLAST which allowed for only one mismatch. In addition, the primer pair for PCR-5A generated a hypothetical product of unexpected size (4,579 bp) using in silico PCR. Assaying for less stringent amplification conditions by increasing the mismatch parameter in the primer BLAST software to two bases however resulted in a different outcome. Possible non-specific amplification products were identified for three additional reactions (PCR-1, PCR-2 and PCR-4) (Data not shown). The expected size of these theoretical products did not however match the observed size of the non-specific products generated following PVL phage typing in the present study.
Figure 3.9: Agarose gel image showing examples of non-specific amplification products following PVL phage typing

Circled on the figure are examples of non-specific PCR amplification products generated following PVL phage typing of clinical isolate TS1 (as described in Section 3.2.3.3).

Expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5: 1,411 or 4,340; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770; PCR-9: 680.
The algorithm for the interpretation of phage type (Ma et al., 2008) was not followed strictly as PCR-5, which is designed to generate specific phage products only from icosahedral-head type PVL phage yielded a product of similar length even for elongated-head type PVL phages. This finding had similarly been reported in a previous study (Otter et al., 2010). Hence, if the PCR results for both icosahedral-head type morphology (PCR-1) and linkage genes (PCR-3) were negative, a positive result for PCR-5 was not taken into consideration in interpretation of the PVL phage typing data.

In total, four of the known PVL-phage types (øPVL, ø108PVL, øSa2USA and øSa2mw) were detected in the present study population (Table 3.8). The majority of isolates (12/19; 63.2%) carried PVL-phage types with icosahedral-head type morphologies. Both of the ST772 isolates in the present study however, failed to generate any PCR product of expected size (Figure 3.10) in all nine PCRs. Though a single faint DNA band was observed in PCR-1, this product was not of the expected size. This complete lack of amplification could perhaps indicate carriage of an uncharacterised PVL phage type by the ST772 isolates.

The phage present in TS21 could also not be fully characterised as no PCR product was generated in the PVL phage specific reactions (i.e. PCR-5 to PCR-9). This result indicates that the PVL phage carried by TS21 does not belong to any of the currently characterised PVL phage types. Unlike the PVL phage carried by ST772 isolates, the PVL phage carried by TS21 could be partially characterised as an elongated-head type PVL phage, based on the positive amplification reactions of PCR-2 and PCR-4.
Table 3.8: Results of PVL-phage typing of the clinical isolates

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate</th>
<th>PCR</th>
<th>PVL Phage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1.</td>
<td>TS 1</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>2.</td>
<td>TS 5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3.</td>
<td>TS 6</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4.</td>
<td>TS 7</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>5.</td>
<td>TS 8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>TS 9</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>7.</td>
<td>TS 12</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>8.</td>
<td>TS 13</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>9.</td>
<td>TS 14</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>10.</td>
<td>TS 15</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>11.</td>
<td>TS 16</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>12.</td>
<td>TS 17</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>13.</td>
<td>TS 18</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>14.</td>
<td>TS 19</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>15.</td>
<td>TS 20</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>16.</td>
<td>TS 21</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>17.</td>
<td>TS 23</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>18.</td>
<td>TS 24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>19.</td>
<td>TS 25</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

DNA amplification reactions to detect the PVL-phage type were set up as described in Section 3.2.2.3. The PVL-phage typing data was interpreted based on the key outlined in Table 3.3 with the modification outlined in Section 3.3.5.

+: Represents the presence of amplification products of expected sizes
+\*: Represents ≈1.4 kb PCR-5 amplification product
+\**: Represents ≈4.3 kb PCR-5 amplification product
-: Represents a lack of amplification product of expected size
***: Elongated head indicates a PVL-phage which could not be fully characterised

Expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5: 4,340 or 1,411; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770.
**Figure 3.10:** Agarose gel image depicting lack of expected amplification products following PVL-phage typing of an ST772 isolate.

The PCR amplification products were generated following PVL phage typing of TS8 (ST772) as described in **Section 3.2.3.3**. The arrow indicates a faint band of unexpected size indicative of a non-specific amplification product.

Expected PCR product sizes (bp): PCR-1: 489 and 569; PCR-2: 468 and 656; PCR-3: 10,497; PCR-4: 9,483; PCR-5: 1,411 or 4,340; PCR-6: 2,238; PCR-7: 4,065; PCR-8: 8,770; PCR-9: 680.
3.3.6 Summary of Results

Overall, the typing results show a clear link between the clonal lineage of an isolate and additional factors such as the antibiotic resistance profiles generated by the NUH, toxin genes detected by HPA, and the presence of specific PVL-phage types as described in the present study (Table 3.9). The exception to this was observed among the ST30 isolates. Despite having identical spa types, these differed greatly in antibiotic resistance, toxin genes present, PVL gene polymorphisms and the PVL phage carried. In contrast, a good correlation between the clonal lineage of an isolate and additional genes was observed with the ST22 isolates. Despite being recovered from a wide variety of skin and soft tissue infections both in the community and the hospital, these isolates showed almost identical antibiotic resistance and toxin profiles and PVL genotypes. They however differed in the spa types they represented.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Date/ Source of Isolation</th>
<th>ST(CC)</th>
<th>C. Laplace</th>
<th>Antibiotic resistances</th>
<th>Virulence</th>
<th>lukS</th>
<th>lukF</th>
<th>Sequence PVL Phage Typing PCR result</th>
<th>PVL-Phage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS 13</td>
<td>17/03/09 (CA)</td>
<td>22(22)</td>
<td>t005</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 14</td>
<td>04/03/09 (HA)</td>
<td>22(22)</td>
<td>t005</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 18</td>
<td>11/11/08 (CA)</td>
<td>22(22)</td>
<td>t005</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 20</td>
<td>03/02/09 (CA)</td>
<td>22(22)</td>
<td>t005</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 24</td>
<td>06/11/08 (CA)</td>
<td>22(22)</td>
<td>t005</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 19</td>
<td>30/12/08 (CA)</td>
<td>22(22)</td>
<td>t310</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 17</td>
<td>04/12/08 (CA)</td>
<td>22(22)</td>
<td>t852</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 6</td>
<td>06/05/09 (HA)</td>
<td>22(22)</td>
<td>t6642</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 9</td>
<td>07/04/09 (HA)</td>
<td>22(22)</td>
<td>t6642</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 23</td>
<td>16/12/08 (CA)</td>
<td>22(22)</td>
<td>t6643</td>
<td>GEN, TMP</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 1</td>
<td>18/03/09 (CA)</td>
<td>1518(152)</td>
<td>t1941</td>
<td>FULLY SENSITIVE</td>
<td>None</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>G T A</td>
</tr>
<tr>
<td>TS 5</td>
<td>23/04/09 (CA)</td>
<td>772(1)</td>
<td>t657</td>
<td>TMP, CIP, ERY</td>
<td>sea, sec, seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 8</td>
<td>08/04/09 (CA)</td>
<td>772(1)</td>
<td>t345</td>
<td>TMP, CIP, ERY</td>
<td>sea, sec, seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 25</td>
<td>05/12/08 (CA)</td>
<td>1(1)</td>
<td>t3342</td>
<td>TMP</td>
<td>sed, seg, sei</td>
<td>A T</td>
<td>T C</td>
<td>T T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 7</td>
<td>04/04/09 (CA)</td>
<td>30(30)</td>
<td>t021</td>
<td>TEMP, DOX</td>
<td>sea, seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C A</td>
<td>A G</td>
</tr>
<tr>
<td>TS 12</td>
<td>18/03/09 (CA)</td>
<td>30(30)</td>
<td>t021</td>
<td>GEN, TEMP, CIP, ERY</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
<tr>
<td>TS 15</td>
<td>12/02/09 (CA)</td>
<td>30(30)</td>
<td>t021</td>
<td>FULLY SENSITIVE</td>
<td>seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C A</td>
<td>A G</td>
</tr>
<tr>
<td>TS 16</td>
<td>11/02/09 (CA)</td>
<td>30(30)</td>
<td>t021</td>
<td>TEMP</td>
<td>sea, seg, sei</td>
<td>G T</td>
<td>T C</td>
<td>C A</td>
<td>A G</td>
</tr>
<tr>
<td>TS 21</td>
<td>11/02/09 (HA)</td>
<td>88(88)</td>
<td>t6769</td>
<td>TEMP</td>
<td>None</td>
<td>G T</td>
<td>T C</td>
<td>C T</td>
<td>A G</td>
</tr>
</tbody>
</table>

Phage (GenBank accession no)
- oSLT (NC_002661)
- o108PVL (AB243556.1)
- oPVL (AB009866.2)
- o2958PVL (AP009363.1)
- oSa2mw (BA000033)
- oSa2USA (CP000730.1)

*Total toxin genes analysed for include: sea, seh, sec, sed, see, seg, seh, sei, sej, tis, eta, eth, and etd. ID, identifier; HA (hospital associated), isolated after 48 h of admission; CA, isolated either at the general practice or accident and emergency department or within 48 h of admission.
*Date of isolation presented as: Day/month/yr.
The *spa* types underlined were described for the first time in the present study.
*GEN: Gentamicin, TEM: Temorphoprim, CIP: Ciprofloxacin, ERY: Erythromycin, DOX: Doxycycline / All isolates were susceptible to clindamycin, rifampicin, linezolid, vancomycin, fusidic acid and teicoplanin
* Asterisks indicate 1,411 bp and 4,340 bp, products respectively. Yellow highlights indicate SNPs
3.4 Discussion

While the molecular epidemiology of CA-MRSA has been explored extensively (Deleo et al., 2010, David and Daum, 2010) and interest in CA-MSSA is on the increase, there exists a dearth of detailed knowledge on PVL-MSSA, particularly in relation to the genetics of the lukSF-PV locus. Despite the fact that initial reports of PVL-positive isolates were described among meticillin susceptible rather than resistant isolates (Lina et al., 1999, Hsu et al., 2005, Prevost et al., 1995a), few studies focused specifically on the typing of these strains. Generally, early typing studies on PVL-MRSA involved identifying their clonal lineages using either PFGE or MLST (Vandenesch et al., 2003, Holmes et al., 2005, Ma et al., 2006, Nimmo et al., 2006, Krziwanek et al., 2007, David et al., 2008, Fang et al., 2008, Goering et al., 2008, Udo et al., 2008, Yu et al., 2008). Some studies went further to associate these with specific spa types (Denis et al., 2005, Tristan et al., 2007, Witte et al., 2007, Bartels et al., 2007, Zhang et al., 2008, Otter and French, 2008, Ellington et al., 2009, Vourli et al., 2009). PVL-MSSA strains were not however entirely left out in the characterisation process. Four studies made mention of PVL-MSSA (Afroz et al., 2008, Karahan et al., 2008, Severin et al., 2008, Orscheln et al., 2009), but none of these carried out a complete typing process on the test isolates.

During the course of the present study, other researchers have also recognised the obvious gaps in knowledge of PVL-MRSA epidemiology and hence, more recent studies have focused on both PVL-MSSA and MRSA strains, often employing both MLST and spa typing. Several studies looked at different populations of PVL-*S. aureus* isolates (Tong et al., 2010, Blanco et al., 2011, Schaumburg et al., 2011,
Brown et al., 2012, Rolo et al., 2012, Zhao et al., 2012), while others simply analysed various sub-groups of PVL-MRSA isolates (Ellington et al., 2010, Holzknecht et al., 2010, Otter et al., 2010, Boakes et al., 2011a, Aschbacher et al., 2012). A few however, focused on PVL-MSSA only (Rasigade et al., 2010, Breurec et al., 2011). Only two of these studies however (Otter et al., 2010, Boakes et al., 2011a), analysed both PVL sequence variation and PVL phage type in relation to strain typing.

PVL-MSSA strains have recently been noted to have an even more diverse genetic background than PVL-MRSA strains and yet the same predominant clonal lineages are represented in both cases (Muttaiyah et al., 2010, Rasigade et al., 2010, Tong et al., 2010). Rather interestingly, in spite of the limited number of isolates analysed in this present study (Table 3.4), a similarly diverse genetic background was detected and the major sequence types associated with PVL-MRSA clones were represented (ST22, ST88, ST30 and ST1). Both ST1 and ST30 are pandemic clones with worldwide distribution. These have routinely been isolated in the UK where they have been reported at one time or the other as one of the major circulating MRSA clones (Ellington et al., 2009, Grundmann et al., 2010, Monecke et al., 2007b, Otter et al., 2009). ST30 is often referred to as the Oceania or South West Pacific clone due to its original isolation from Australia, New Zealand and Western Samoa. In contrast, ST1, a USA400 PFGE related clone was initially described in the USA, though currently its prevalence is in decline (Vandenesch et al., 2003, Tristan et al., 2007).

The PVL-MRSA ST22 isolates, though related to the pandemic HA-MRSA ST22 EMRSA-15 clone are in themselves not pandemic but rather have been described
predominantly in Europe and the UK (Ghebremedhin et al., 2005, Holmes et al., 2005, Krziwanek et al., 2007, Monecke et al., 2007b). In the UK, while the majority of PVL-MRSA circulating clones belong to either ST8, ST30 or ST80 (Holmes et al., 2005, Ellington et al., 2009, Otter and French, 2008), isolates belonging to ST22, the predominant sequence type found in the present study (Table 3.9), also constitute a significant burden of infection (Ellington et al., 2009). Of the other clones identified in the present study, ST772 was first described in Asia (Afroz et al., 2008, D'Souza et al., 2010), and has an MLST allelic profile of 1-1-1-1-22-1-1. This is a single locus variant of ST1 and belongs to the same clonal complex, CC1. Recently, an increase in the number of ST772 isolates was noted (Ellington et al., 2010) in the UK. These isolates are known as the ‘Bengal Bay clone’ in England, due to an association with importation from India and Bangladesh. Our test isolates (Table 3.9) therefore appear to reflect the current distribution of S. aureus strains in England.

The new ST identified in the present study, ST1518 with an MLST allelic profile of 46-75-207-44-13-68-60, is a single locus variant of ST152 (46-75-49-44-13-68-60). ST152 is a multilocus ST which is not commonly found in the UK, but has been reported in several European and African countries. In addition, it is thought to be associated with travel, specifically of immigrants from ex-Yugoslavian countries (Blanc et al., 2007, Monecke et al., 2007a, Berglund et al., 2005, Okon et al., 2009). ST152 is interesting because all isolates so far identified are PVL-positive, and it has generally been associated with carriage and skin and soft tissue infections (Muller-Premru et al., 2005, Perez-Roth et al., 2010, Ruimy et al., 2008), rather than severe life threatening infection. In contrast, the TS1 Nottingham strain which belongs to CC152 was isolated from a patient who died from necrotising pneumonia.
Where MLST provides a snapshot of global epidemiology by looking at strain relatedness based on long term changes, the information it provides is sometimes inadequate as highlighted in a 2009 study (Larsen et al., 2009a). In this study, two rather distinct clones presented with identical PFGE and MLST profiles (USA300, ST8). These clones were however associated with clearly different clinical profiles, varying in age of patients infected and location of acquisition. The isolates also varied in the type of MGEs carried and antibiotic susceptibility profiles (Larsen et al., 2009a). The two clones corresponded to two different spa types, t008 and its single repeat variant t024. The t008 clone was found to be PVL-positive while the t024 clone was PVL-negative. This identical delineation of spa type and PVL encoding genes between t008 and t024 strains had also previously been reported during a study characterising a group of isolates from Denmark (Bartels et al., 2007).

Another group of isolates (the ST22) further highlight the need for spa typing as an essential adjunct to the characterisation of S. aureus. Over the years, various studies analysing different collections of S. aureus isolates described an apparent association between PVL and the ST22, EMRSA-15 strains (Udo et al., 2006, Coombs et al., 2009, Nadig et al., 2010). These reports of the presence of PVL encoding genes in a highly successful hospital acquired clone raised concerns at the potential implications if this phenomenon became widespread, with respect to increase in virulence and severity of infection. In these three studies, isolates were described as EMRSA-15 based on PFGE pattern, phage typing, production of SEC, non-production of urease and resistance to erythromycin and/or ciprofloxacin (Udo et al., 2006), or MLST and type of SCCmec element (Coombs et al., 2009), or MLST and spa type (Nadig et al.,
Two other studies which described an association between the PVL toxin and ST22 isolates more cautiously referred to these isolates as EMRSA-15 related clones (Holmes et al., 2005, Goering et al., 2008) presumably based on the meticillin susceptible status of the strains.

A closer analysis of the literature shows that like the t008 and t024 spa types of ST8, the presence of the lukSF-PV locus in ST22 isolates varies depending on the spa type. Among the ST22 isolates, PVL encoding genes have only been rarely described in the two spa types which are more commonly associated with EMRSA-15 (t032 and t022) (Aires-de-Sousa et al., 2008, Boakes et al., 2011b, Otter and French, 2008). Therefore, despite reports of PVL-positive ST22 isolates, PVL encoding genes are rare in EMRSA-15 associated spa types. This was demonstrated in a recent study exploring the molecular diversity of CC22 PVL-MRSA strains in the UK (Boakes et al., 2011b). None of the 47 test isolates belonged to either spa type t032 or t022 but represented among these strains were spa types t005, t852, t1516, t2816, t1760, t849 and t1790. This absence of spa types t032 and t022 was similarly maintained within our population of PVL-MSSA ST22 strains (Table 3.9). spa typing analyses the variable region of the spa gene which is highly polymorphic. Unlike both PFGE and MLST typing, spa typing generally reflects local, more recent, variations. It has a higher discriminatory power than MLST, as clearly demonstrated in the present study whereby five different spa types (t005, t310, t852, t6642 and t6643) were represented within ST22 alone.

Furthermore, in this present study, the ST22 isolates analysed were genetically similar in susceptibility and toxin profile as well as spa type to the recently described ST22
PVL-positive MRSA UK strains (Ellington et al., 2010), rather than to the predominant HA-MRSA ST22 EMRSA-15 clone (O'Neil et al., 2001, Wolter et al., 2008, Grundmann et al., 2010). This is in agreement with the observation by Boakes and colleagues (Boakes et al., 2011b), who further went on to assert that PVL-ST22 strains in the UK are evolving independently rather than by EMRSA-15 acquiring the PVL encoding genes.

The genes encoding the Panton Valentine leucocidin may be carried by several temperate bacteriophages, of which at least five have been described in the literature (Kaneko et al., 1998, Kaneko et al., 1997, Narita et al., 2001, Baba et al., 2002, Diep et al., 2006b, Ma et al., 2006, Ma et al., 2008). Notwithstanding the differences in genetic background of the genes encoding the PVL toxin, the sequence of both lukF and lukS genes is relatively conserved with 12 major single nucleotide polymorphisms (SNPs) observed in total (Takano et al., 2008b). This previously described highly conserved nature of the genes encoding the PVL toxin (O'Hara et al., 2008) was also observed in the present study (Table 3.7). A total of 7 SNPs were detected and no additional polymorphisms described. In agreement with reports made from previous studies on lukSF-PV locus polymorphism (Otter et al., 2010, Takano et al., 2008b), PVL typing of select isolates showed that specific clonal lineages harboured the same PVL encoding gene sequences despite the variations in spa types within these clones (Table 3.9). A comparison of the lukSF-PV gene sequences from the present study with the current literature revealed that while specific clones appear to be associated with a specific sequence, the same lukSF-PV gene sequence variant may be carried by more than one lineage (Table 3.9).
These sequence polymorphisms have acted as the basis for the haplotypes previously defined by O’Hara and colleagues (O’Hara et al., 2008). Both geographical and clonal bias has been shown to exist in the distribution of the two haplotypes defined by these polymorphisms. The R haplotype traditionally occurs within a limited number of lineages predominantly in the USA and more recently Australia (O’Hara et al., 2008, Tong et al., 2010). Due to previous reports of isolates carrying R haplotypes being found in the UK in significant numbers (Otter et al., 2010, Boakes et al., 2011a), the low incidence of R haplotypes in this small study population may reflect more on the meticillin susceptibility of the isolates rather than geographical location. No significant functional and structural differences are thought to result from the arginine (R) to histidine (H) mutation present in this haplotype (Besseyre des Horts et al., 2010, Berglund et al., 2008b), however it is of interest that the only R haplotype found in this present study was linked with a fatal case of necrotizing pneumonia in a healthy adult. By 2011, this appeared to be the first reported case of an isolate carrying the R isoform described outside the previously noted clonal complexes (CC1, CC5, CC8 and ST93). Two recent publications have however reported this isoform in three isolates which belonged to ST30, ST121 and CC377 respectively (Li et al., 2012, Brown et al., 2012). This increased detection of the R isoform in novel backgrounds may simply reflect the recent increase in PVL typing of S. aureus strains.

One evolutionary pathway postulated for the development of CA-MRSA is the acquisition of PVL encoding genes prior to that of the SCCmec element. To lend credence to this theory, one would expect to find strains of both PVL-MSSA and PVL-MRSA which are more or less similar but simply differ in the presence of the mec gene and perhaps other resistance genes. Of the 11 unique MLST/spa clones
described in the present study, 4 (36.4%) of these were described for the first time (ST22/t6642, ST22/t6643, ST1518/t1941 and ST88/t6769), and one other had only just been described in 2009 (ST1/t3342). Of the 6 remaining profiles, limited typing data was available in the literature with regards to ST22/t310 and ST30/t021 isolates, therefore, there was no basis for comparison. A single study identified two fully sensitive MSSA ST30/t021 strains from Ireland with the same virulence profile as found in TS15 in the present study (Ellington et al., 2007). ST22/t005, ST22/t852, ST772/t345 and ST772/t657 were however found to possess spa and toxin profiles identical to clones of circulating CA-MRSA in the UK (Ellington et al., 2009, Ellington et al., 2010). These varied only in susceptibility profile, with the CA-MRSA strains generally presenting a more resistant profile, thereby lending support to the evolution of PVL-MRSA from PVL-MSSA.

The number of MSSA isolates in the present study having identical PVL isoforms as their previously described MRSA counterparts also lends further credence to the hypothesis of PVL-MRSA evolving from PVL-MSSA (Berglund et al., 2008b, Boakes et al., 2011a, Dumitrescu et al., 2008b). This process may however not be limited to a single evolutionary event. The detection of several PVL phage types observed in the ST30 lineage in the present study, and described by Boakes and colleagues in two separate studies (Boakes et al., 2011b, Boakes et al., 2011a), hints at the possibility of multiple phage acquisitions and hence evolutionary events, though this remains to be verified.

While the data generated in the present study, appears to favour evolution of PVL-MRSA from PVL-MSSA, it is important to note that in contrast, several PVL-negative
CA-MRSA clones (as described by MLST and spa data) have been reported, which appear to have similar profiles to PVL-MRSA strains (Bartels et al., 2007, Otter and French, 2008, Zhang et al., 2008, Rivero-Perez et al., 2012). Otter and French described two MRSA clone groups (ST1/t127 and ST8/t008) which contained both PVL-positive and PVL-negative strains and hence an argument could be made for the acquisition of the mec genes prior to the genes encoding the PVL toxin. Antibiogram data in these studies though were not directly linked with each specific isolate thereby preventing any further comparisons or conclusions to be drawn. Another study highlights the possibility of acquisition of mec genes prior to genes encoding the PVL toxin (Zhang et al., 2008). In this study, similar PVL-positive and PVL-negative variants of USA400/t128 were detected. These strains however, simply differed in the carriage of fnbB and cnaA virulence genes. These findings may therefore suggest that evolution of PVL-MRSA can take place by several pathways. Thus, further research will be required to ascertain if all clones evolve by either pathway, or if several clones favour a certain pattern of evolution.

Despite the increase in discriminatory power of the initial PVL-phage typing method with the recent description of additional primers enabling øSa2USA detection (Boakes et al., 2011a), the inability of this technique to detect the ST772 PVL phages highlight its limitations and the need for constant updating to encompass evolving novel phages.

In conclusion, due to the gentamicin/trimethoprim resistance bias in isolate selection, the present study provides one of the most comprehensive in-depth analysis of PVL-MSSA ST22 isolates incorporating data on the PVL encoding genes to date. It also highlights the important role PVL-MSSA strains generally play as reservoirs for the
emergence of PVL-MRSA strains due to their direct evolutionary links suggested by the typing results. The data emphasise the need for increased surveillance which may form the basis for intervention strategies against PVL-MSSA to help curb the emergence and clonal expansion of PVL-MRSA.
Chapter Four

4 Development and validation of a high resolution melt based method for the rapid typing/identification of PVL-positive strains of \textit{S. aureus}

4.1 Introduction

Despite the wealth of information provided by the typing systems employed in the initial part of this research, these current typing methods for \textit{S. aureus} are often cumbersome, expensive and time consuming. This can be illustrated by an analysis of the process for the sequence based typing systems (Figure 4.1). A number of the typing systems are simply based on sequence comparison. This involves the need to ascertain the presence, integrity and purity of a PCR product by visualisation on an agarose gel following both the amplification and purification processes. All these steps combine to increase the complexity and timescale of the technique before the actual sequencing take place. As the MLST procedure involves the analysis of seven unique loci per isolate, significant manpower and time resources are consequently involved in obtaining the essential typing information, highlighting the need for an alternative.
Figure 4.1: Process flowchart for sequence based typing systems

The approximate timings of each stage in the process of the sequence based typing systems are indicated in brackets.
4.1.1 **High Resolution Melt (HRM)**

The recently described high resolution melt (HRM) technique is a real time PCR based method. This method eliminates the need for product detection via the traditional post amplification processes such as agarose gel electrophoresis. Furthermore it enables PCR amplicon differentiation based on variations in DNA melt characteristics (Gundry et al., 2003, Wittwer et al., 2003). Since its development, HRM has found wide application mainly in SNP detection and as a gene scanning tool. A few studies have employed this technique as a bacterial typing tool based on variations in a single gene. These variations in application are clearly highlighted by a comparison of two publications (Price et al., 2007, Schwartz et al., 2009). Whereas one study applied HRM to genotyping of *Campylobacter jejuni* based on size variations in a class of short sequence repeats (CRISPRs) (Price et al., 2007), a second study (Schwartz et al., 2009) identified different *Mycoplasma pneumoniae* isolates based on SNPs in a single constant length gene fragment.

The ≈1.9 kb *lukSF-PV* locus which encodes both LukF and LukS subunits of the Panton-Valentine leucocidin is a highly conserved region in which 12 major SNPs have been described to date ([Section 1.6.3](#)) (Takano et al., 2008b, Berglund et al., 2008b, Dumitrescu et al., 2008b). Only four of these SNPs are however commonly found within PVL-positive isolates. These occur at positions 527 and 663 of *lukS* and 1396 and 1729 of *lukF* and represent two major allelic variations per gene. A\_527/G\_663 and G\_527/T\_663 in the *lukS* gene and G\_1396/A\_1729 and A\_1396/A\_1729 in the *lukF* gene ([Table 4.1](#)). Furthermore, the *lukS* locus is unique in the fact that the two predominant
variants (A_{527}/G_{663} and G_{527}/T_{663}) potentially balance each other out thermodynamically.

These four SNPs serve as the basis for the designation of previously described H and R isoforms of lukSF-PV (O'Hara et al., 2008). These allelic variations in both lukF and lukS genes have been shown to be specific (though not exclusive) to the different S. aureus clones as defined by STs (Table 1.4). The potential therefore exists for the use of these variations in HRM based genotyping of PVL-positive S. aureus isolates.

The use of HRM in detecting all previously described classes of single base pair changes within a specific amplicon has been well documented (Krypuy et al., 2006, Toi and Dwyer, 2008, Sinthuwiwat et al., 2008, Norambuena et al., 2009). Shorter amplicons ranging in size from 45 – 190 bp were generally employed to improve result specificity. Only a limited number of studies have however made mention of the detection of a double mutation within an amplicon (Pietzka et al., 2009, van der Stoep et al., 2009, De Leeneer et al., 2008).
Table 4.1: The four major SNP combinations used to define PVL isoforms

<table>
<thead>
<tr>
<th>lukSF-PV sequence variants</th>
<th>527*</th>
<th>663*</th>
<th>1396*</th>
<th>1729*</th>
<th>Main STs**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>22, 59, 88</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>30, 50, 80, 121, 772</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>8, 93, 1, 5</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>1</td>
</tr>
</tbody>
</table>

*These values indicate nucleotide positions of the allelic variations within the lukSF-PV locus
**Specific STs associated with each variant combination as compiled from (Takano et al., 2008b, Berglund et al., 2008b, Otter et al., 2010, Wolter et al., 2007)

The four major SNPs shown are as described by O’Hara et al. (2008)
The present study therefore set out to analyse the ability of HRM to differentiate between allelic variations including those which potentially balance each other out thermodynamically (as seen in the lukS locus). This technique was then explored as a possible tool for rapid molecular characterisation of PVL-positive S. aureus isolates based on the specific allelic variations present in the lukSF-PV locus. If successful, this method could significantly reduce turnaround typing time and improve outbreak investigation and epidemiological studies.

Studies presented in this chapter aimed to answer the following questions.

1. Could the HRM technique be used to differentiate between both A\textsubscript{527}/G\textsubscript{663} and G\textsubscript{527}/T\textsubscript{663} variants of the lukS gene as well as the lukF gene variants?
2. Could the HRM technique be used to class isolates of differing spa types as either same or different?
3. Could this technique then be applied as a genotyping tool for a group of blinded isolates?
4. Are the results generated by this technique consistent and accurate?
4.2 Materials and Methods

4.2.1 Primers

To explore the use of HRM analysis in the rapid typing of PVL-positive *S. aureus* isolates at both the *lukSF-PV* and *spa* loci, three sets of primers were used (Table 3.1). The *lukSF-PV* primers were designed to capture allelic variations at positions 527 and 663 of the *lukS* gene and at positions 1396 and 1729 of the *lukF* gene, thereby generating PCR products of 272 bp and 447 bp in length respectively. Previously described primers were used for the *spa* locus (Aires-de-Sousa et al., 2006).

4.2.2 Standard S. aureus Isolates

Four standard NARSA (Network on Antimicrobial Resistance in *S. aureus*) isolates NRS123, NRS157, NRS158 and NRS162 with known allelic variations of G527/T663-A1396/A1729; A527/G663-G1396/A1729; A527/G663-G1396/A1729; and G527/T663-G1396/A1729 respectively at positions of the *lukSF-PV* locus indicated above were used in the development of the *lukSF-PV* HRM (Table 4.2). Development of the *spa* HRM was similarly carried out using five PVL-positive *S. aureus* isolates with known *spa* types (Table 4.3), from our own strain collection.
Table 4.2 Allelic variation of standard isolates used for the development of the **lukSF-PV** HRM assay

<table>
<thead>
<tr>
<th>S/No</th>
<th>Strain ID*</th>
<th>Allelic Variation**&lt;br&gt;lukS – lukF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>NRS123</td>
<td>G_{527}/T_{663} A_{1396}/A_{1729}</td>
</tr>
<tr>
<td>2.</td>
<td>NRS157</td>
<td>A_{527}/G_{663} G_{1396}/A_{1729}</td>
</tr>
<tr>
<td>3.</td>
<td>NRS158</td>
<td>A_{527}/G_{663} G_{1396}/A_{1729}</td>
</tr>
<tr>
<td>4.</td>
<td>NRS162</td>
<td>G_{527}/T_{663} G_{1396}/A_{1729}</td>
</tr>
</tbody>
</table>

*Standard isolates were obtained from the Network on Antimicrobial Resistance in *S. aureus* (NARSA) collection.

**Allelic variations are as described in Table 4.1 representing variations at positions 527, 663, 1396 and 1729 of the **lukSF-PV** locus.
<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID</th>
<th>spa type</th>
<th>Repeat Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TS6</td>
<td>t6642</td>
<td>26-23-13-23-31-05-22-17-25-17-25-114</td>
</tr>
<tr>
<td>2</td>
<td>TS9</td>
<td>t6642</td>
<td>26-23-13-23-31-05-22-17-25-17-25-114</td>
</tr>
<tr>
<td>3</td>
<td>TS14</td>
<td>t005**</td>
<td>26-23-13-23-31-05-17-25-17-25-16-28</td>
</tr>
<tr>
<td>4</td>
<td>TS18</td>
<td>t005</td>
<td>26-23-13-23-31-05-17-25-17-25-16-28</td>
</tr>
<tr>
<td>5</td>
<td>TS17</td>
<td>t852**</td>
<td>07-23-13-23-31-05-17-25-17-25-16-28</td>
</tr>
</tbody>
</table>

The isolates chosen represented spa types generating amplicons of identical lengths (as indicated by number of repeats) but varying in nucleotide composition (Isolates were obtained from the NUH Trust between 2008 and 2009).

** spa types t005 and t852 differ at only a single base pair.
4.2.3 HRM

Using the Roche LC480 High resolution mastermix, samples were set up for PCR amplification in a 25 µl final volume containing 1 × MasterMix, 1 µl template DNA, 300 nM each primer and 2.5 mM MgCl₂ on a Roche LightCycler® 480 Real-Time PCR System, using a 96-well plate. The amplification protocol involved a 5 min hold at 95°C followed by 45 amplification cycles comprising of 95°C denaturation for 10 s, annealing at 60°C for 10 s and elongation at 72°C for 10 s. The HRM reactions were carried out in triplicate for each isolate tested.

After amplification, high resolution melting (Section 1.7.4) was carried out on the PCR amplicons generated. The first part of the melting process involved a brief denaturation of samples at 95°C for 1 min and rapid annealing. Following this, the temperature was again slowly increased and HRM curve data obtained from 65°C to 97°C with 5 fluorescence acquisitions per degree centigrade (1°C).

Finally, HRM analysis involving normalisation, temperature shift and difference curve generation was carried out on the melt data using the automatic settings of the Gene Scanning Software Version 1.5.0 (Roche Diagnostics) except were indicated. Normalisation involves the conversion of raw fluorescence data of the active melt region into a relative fluorescence signal ranging from 0% (≈ lowest fluorescence value) to 100% (≈ highest fluorescence value). Conversely, temperature shift involves shifting on the X – axis to show up the relative rather than absolute temperature difference between the curves (Reja et al., Wittwer et al., 2003). Difference curves are
generated by subtracting the normalised and temperature shifted curves from that of a reference ‘base curve’. This process makes it easier to visualise small variations between amplicons.
4.3 Results

4.3.1 *lukSF-PV* locus HRM Analysis

An analysis of the melt profile of the four NARSA isolates at the 272 bp *lukS* locus revealed four distinct but similar melt profiles and peaks (Figure 4.2 A), with melting temperatures ($T_m$) ranging from 76.25°C to 76.68°C. Subsequent HRM analysis involving normalisation and temperature shift using the Roche Gene Scanning Software version 1.5.0, classed the four standard isolates into two distinct groups (Figure 4.2 B). The variations within these two groups were then more clearly represented by the generation of a difference curve using the Roche Gene Scanning Software (Figure 4.2 C) using the automatic settings.

The melt profile of the 447 bp *lukF* product differed from that of the 272 bp *lukS* locus with the product melting in two steps (Figure 4.2 D). Unlike the HRM profiles produced following analysis of the *lukS* locus, a visual inspection of the *lukF* HRM profiles did not provide clear cut demarcation into distinct groups. This inability to visually separate the *lukF* profiles into groups did however have an effect on the final outcome of the analysis process. The gene scanning software was able to automatically cluster the *lukF* amplicons into two distinct groups after normalisation and temperature shift.
Figure 4.2: Representation of different melt curves generated during the HRM analysis of four *S. aureus* NARSA isolates.

Panel A presents the initial melt profiles of the four standard isolates at the *lukS* locus showing no obvious clear groupings. Panel B presents the normalised and temperature shifted melt profiles of the same four standard isolates showing a clear demarcation of the two allelic variants represented within the four isolates into two groups. Panel C presents the difference curves of the four standard isolates at the *lukS* locus highlighting the difference in high resolution melt profile between the variants. Panel D presents the *lukF* normalised and temperature shifted melt profile showing multiple melt domains which were assigned into two groups by the HRM software as represented by the two colours.
The four different allelic variations $\text{G}_527/\text{T}_{663}$; $\text{A}_527/\text{G}_{663}$; $\text{A}_1396/\text{A}_{1729}$ and $\text{G}_1396/\text{A}_{1729}$ present in these NARSA isolates were hence designated S1, S2, F1 and F2 respectively.

The next step in the process was to explore the ability of the HRM software to accurately predict the genotypes of unknown isolates based on known standard genotypes. To do this, melt profiles of NRS123 and NRS157 with genotypes S1 and S2 respectively, were used as positive internal standard controls. Based on the genotypes of these two isolates, the gene scanning software was able to accurately class isolates NRS158 and NRS162 which were used as the unknowns, as S2 and S1 (Figure 4.3). This result therefore demonstrated the ability of the technique to determine the genotypes of unknown isolates based on HRM profiles of standard isolates.
Figure 4.3: Automatic grouping of NRS158 and NRS162 as S2 and S1 variants (circled) respectively

Grouping of standard strains NRS158 and NRS162 following HRM analysis was automatically performed by the Gene Scanning Software Version 1.5.0 based on the data obtained from the specified internal control isolates NRS123 and NRS157.
4.3.1.1 **Pilot study to characterise a subset of isolates using HRM analysis**

Following the demonstration of the ability of the HRM technique to distinguish between both *lukS* and *lukF* variants, a test run was carried out to further explore the potential of the technique as a genotyping tool. The result of this analysis using 20 previously characterised PVL-positive *S. aureus* test isolates, highlighted the potential of this method (**Table 4.4**).

HRM analysis of the *lukS* locus resulted in 100% typeability. All isolates were assigned to either group S1 or S2. At the *lukF* locus, 95% (19/20) of isolates could be assigned to one of the two previously described groups i.e. F1 or F2. A single isolate (TS1) was classed as ‘unknown’ (FU). Isolate TS1 differed from both the F1 and F2 control genotypes by having an allelic profile of $A_{1396}/G_{1729}$, further highlighting the specificity of the HRM assay.

A total of three profiles (S1:FU, S2:F1, and S2:F2) were therefore generated for the twenty isolates (**Table 4.4**) corresponding to the six STs present within this collection.
Table 4.4: HRM profiles of isolates generated following HRM based typing

<table>
<thead>
<tr>
<th>S/NO</th>
<th>HRM Profile*</th>
<th>Test Isolate</th>
<th>ST**</th>
<th>lukS 527</th>
<th>663</th>
<th>lukF 1396</th>
<th>1729</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>S1:FU</td>
<td>TS1</td>
<td>1518</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>2.</td>
<td>S2:F1</td>
<td>TS25</td>
<td>1</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>3.</td>
<td>&quot;</td>
<td>TS7</td>
<td>30</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>4.</td>
<td>&quot;</td>
<td>TS15</td>
<td>30</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>5.</td>
<td>&quot;</td>
<td>TS16</td>
<td>30</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>6.</td>
<td>&quot;</td>
<td>TS2</td>
<td>772</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>7.</td>
<td>&quot;</td>
<td>TS5</td>
<td>772</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>8.</td>
<td>&quot;</td>
<td>TS8</td>
<td>772</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>9.</td>
<td>S2:F2</td>
<td>TS6</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>10.</td>
<td>&quot;</td>
<td>TS9</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>11.</td>
<td>&quot;</td>
<td>TS13</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>12.</td>
<td>&quot;</td>
<td>TS14</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>13.</td>
<td>&quot;</td>
<td>TS18</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>14.</td>
<td>&quot;</td>
<td>TS17</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>15.</td>
<td>&quot;</td>
<td>TS19</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>16.</td>
<td>&quot;</td>
<td>TS20</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>17.</td>
<td>&quot;</td>
<td>TS23</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>18.</td>
<td>&quot;</td>
<td>TS24</td>
<td>22</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>19.</td>
<td>&quot;</td>
<td>TS12</td>
<td>30</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>20.</td>
<td>&quot;</td>
<td>TS21</td>
<td>88</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
</tbody>
</table>

*The HRM profile was generated for each isolate by combining HRM data from both the _lukS_ and _lukF_ loci.

**MLST sequence type.
4.3.2 spa locus HRM analysis

Despite the promising typing results obtained in Section 4.3.1.1, the level of discrimination produced was 50% lower than that provided by MLST. HRM analysis of DNA fragments of the *lukSF-PV* locus generated three genotypes as opposed to the six STs represented within the test population (Table 4.4). Therefore, the next step in the present study was to explore the use of HRM in genotyping *S. aureus* isolates based on variations in *spa* types. This step was taken in a bid to improve the potential utility of HRM as a typing tool by possibly combining data from the *lukSF-PV* and *spa* loci.

4.3.2.1 Test strain characteristics

Five isolates representing three different *spa* types were used for the development of the *spa* HRM technique (Table 4.3). These *spa* types were chosen to demonstrate the ability of the technique to differentiate *spa* amplicons with identical lengths but variable nucleotide composition. The amplicons from all three *spa* types represented were 288 bp in length (Table 4.3). Additionally, isolate TS17 (t852) was included in the analysis to explore the ability of the HRM technique to differentiate between *spa* types t852 and t005 which differ by only a single base pair.
4.3.2.2 HRM analysis of isolates of differing spa type

Initial analysis of the spa HRM data using the automatic settings of the Roche Gene Scanning Software, generated complex shaped difference curves with more than one peak (indicated by the arrows in Figure 4.4 A). These complex shaped plots negatively affected discrimination as crossing over and overlapping of the difference curves (indicated by the circle) limited visual interpretation (Figure 4.4 A). This limitation was overcome by carrying out HRM analysis with a temperature shift of zero (0°C) on the data. The process resulted in single peaks and no overlaps (Figure 4.4 B). Hence all further analysis was carried out with a temperature shift of zero.
Figure 4.4: Normalised and temperature shifted difference curves of five S. aureus isolates representing three different spa types using temperature shifts of 5°C or 0°C.

The HRM profiles which were generated as described in Section 4.2.3 show the effect of temperature shift on HRM analysis of the spa locus. Panel A presents the complex shaped difference curves generated with the automatic temperature shift of 5°C in place. Panel B presents the clearly resolved difference curves when a temperature shift of 0°C was used.

Colour code: Blue = TS17 (t852); Yellow = TS14 (t005); Orange = TS18 (t005); Pink = TS6 (t6642); Red = TS9 (t6642). The spa types are shown in bracket
4.3.2.3 *spa* HRM Data Interpretation

Following HRM analysis, assignment of the different *spa* types as same or different was carried out using previously defined criteria (Stephens et al., 2008). Isolates generating similar shaped difference curves falling within 4 relative signal difference (RSD) units of each other were grouped as identical. RSD is the unit of measurement generated when a difference curve is created.

Both t6642 isolates (TS6 and TS9) consistently (over 5 repeat experiments) produced similar difference curves with RSD values which differed by <4 (Figure 4.5). This enabled both isolates to be grouped as same. Interpretation of results for the t005 isolates (TS14 and TS18) however proved more problematic. While generating similar shaped difference curves, these consistently had RSD values which differed by >4 (≈18). In contrast, TS17 (t852) and TS14 (t005) produced more similar curves with differences in RSD values closer to 4 (≈5). This similarity was highlighted by the automatic classification of these two isolates as same by the gene scanning software (Figure 4.6).
Figure 4.5: Normalised and zero temperature shifted difference curves of five *S. aureus* test isolates representing three different *spa* types.

The HRM profiles were generated as described in Section 4.2.3 with a 0°C temperature shift. The figure shows similar HRM profiles generated by t6642 isolates (TS6 and TS9) with a difference in RSD values of <4. Additionally, t005 isolates (TS14 and TS18) generated similar shaped HRM profiles but a difference in RSD value of ≈18. t852 (TS17) generated a HRM profile similar to TS14 with a difference in RSD value of ≈5.

Colour code: Blue = TS17 (t852); Yellow = TS14 (t005); Orange = TS18 (t005); Pink = TS6 (t6642); Red = TS9 (t6642). The *spa* types are shown in bracket.
Figure 4.6: Automatic classification of normalised and zero temperature shifted difference curves of five *S. aureus* test isolates into three groups.

The HRM profiles were generated as described in Section 4.2.3 with a 0°C temperature shift. The figure shows the automatic classification of HRM profiles of both t6642 isolates (in blue) as same, misclassification of the t005s (in red and green respectively) as different and t852 (T17) wrongly classed as identical to t005 (TS14). Each individual group as defined by the gene scanning software is represented by a different colour.
In order to explore the discrepancies noted between the t005 isolates in Figure 4.6, further HRM analysis of a total of five isolates of *spa* type t005, was carried out. This analysis resulted rather interestingly in their classification into more than one group. While all the difference curves corresponding to the five t005 isolates were similar in shape, they differed greatly in RSD. Of the five isolates, TS13, TS14 and TS18, produced difference curves which clustered together (Figure 4.7), with a difference in RSD values of >4 (≈38.2, 36.2 and 34.5 respectively). HRM data for these three isolates was obtained from the first 6 columns of the 96-well plate. TS20 and TS24 produced similar difference curves. These curves however had RSD values which differed by >4 (≈24.2 and 11.3 respectively). As HRM data for these two isolates was obtained from columns 8 and 10 of the 96-well plate, this suggested a possible effect of well position on result outcome.

This possible effect of well position on assay outcome was further highlighted in the case of TS14 and TS18, which produced HRM profiles with a difference in RSD values <4. This result was in contrast to the results of the initial assay (Figure 4.4 B) whereby a difference in RSD value of ≈18 was noted between the isolates. The only variable between the two assays was the well position used in the microtitre plate. In the initial assay, TS14 and TS18 isolates were analysed in columns 6 and 10 respectively while they were analysed in columns 4 and 7 respectively in the second assay.
Figure 4.7: High resolution melt profile of five *S. aureus* isolates with identical t005 *spa* types

The HRM profiles were generated as described in Section 4.2.3 with a 0°C temperature shift. The figure shows HRM profiles of isolates to the right of the plate (colour coded blue and orange) presenting with differences in RSD values >4 and TS14 and TS18 formerly classified as different presenting difference curves clustered together. Inset represents the different colour codes associated with each plate column.

Colour code: Red = TS13; Pink = TS14; Yellow = TS18; Blue = TS20; Orange = TS24
4.3.2.4 Effect of well position on results of HRM analysis

In order to further explore the possible effect of well position on HRM profiles, two randomly chosen strains, TS6 and TS9 of spa type t6642 were used. Twelve aliquots of a large scale PCR “mastermix” created for each strain were dispensed into wells at various positions of a 96-well plate (Figure 4.8) and PCR amplification and HRM analysis performed.

Figure 4.8: Aliquot positions used to assay for effects of well position on HRM profiles.

The aliquot positions for TS6 are highlighted in red while the aliquot positions for TS9 are highlighted in black.
The amplification efficiency of the different aliquots of both TS6 and TS9 appeared to vary across the plate, resulting in variable amplification curves (Figure 4.9). These amplification curves which represent the fluorescence generated per amplification cycle, differed in the fluorescence value at which the curves entered the non-exponential plateau phase. For most reactions of both isolates (7 of 12), a plateau phase was reached at fluorescence units ranging from 64 – 74. Of the 5 reactions with a plateau phase corresponding to fluorescence units over 74, four occurred in columns 1 and 12 of the 96-well plate.

Similarly, well position appeared to impact on the outcome of HRM analysis (Figure 4.10). Samples in columns 11 and 12 of the 96-well plate gave RSD values which differed by >4. This resulted in amplicons in these columns being grouped separately despite the similar shaped difference curves (Figure 4.10). This analysis provided a strong indication that the observed well/column based variability could be a reflection of temperature variability across the plate.

In order to further explore this hypothesis, the automatic temperature shift function of the gene scanning software was restored (Figure 4.11). Temperature shift, one of three functions of the gene scanning software, is designed to correct for temperature variation across the plate. While restoring the automatic temperature shift function at this stage resulted in more complex shaped difference curves, it did result in a reduction in differences in RSD values between replicate samples analysed at different positions on the plate (Figure 4.11). A position dependent clustering (of samples) was however still maintained, highlighting the need to take into consideration sample position in the interpretation of spa HRM data generated by the Roche LC480.
Figure 4.9: Variable amplification profiles generated for aliquots of single master mixes for TS6 and TS9 analysed at different well positions of the 96-well plate.

The amplification curves were generated during amplification as described in Section 4.3.2. The figure shows differences in fluorescence units corresponding to non-exponential plateau phase with four out of the five samples producing fluorescence units >74 located in columns 1 and 12. The inset represents the different colour code associated with each plate column Colour code: Blue = Column 1; Red = Column 2; Yellow = Column 6; Purple = Column 11; Green = Column 12 (Aliquots positions of TS6 are highlighted in red and TS9 in black).
Figure 4.10: Normalised and zero temperature shifted difference curves of a single sample (TS6) analysed at different positions of the 96-well plate.

The HRM profiles were generated as described in **Section 4.2.3** with a 0°C temperature shift. The figure shows HRM profiles of a single master mix dispensed at varying well positions clustering into two distinct groups dependent on sample location on the plate (as indicated by colour code). The inset represents the different colour code associated with each plate column Colour code: Blue = Column 1; Red = Column 2; Yellow = Column 6; Purple = Column 11; Green = Column 12 (Aliquots positions of TS6 are highlighted in red and TS9 in black)
Figure 4.11: Normalised and automatic temperature shifted difference curve clustering for aliquots of single master mixes for TS6 and TS9 analysed at different well positions of the 96-well plate.

The HRM profiles were generated as described in Section 4.2.3 with a 0°C temperature shift. The figure shows HRM profiles of a single master mix dispensed at varying well positions showing a reduction in RSD and less defined difference curve grouping dependent on sample location on the plate (as indicated by colour code). The inset represents the different colour code associated with each plate column Colour code: Blue = Column 1; Red = Column 2; Yellow = Column 6; Purple = Column 11; Green = Column 12 (Aliquots positions of TS6 are highlighted in red and TS9 in black). The arrows indicate RSD values showing differences in values slightly greater than 4 between aliquots in column 11 and the aliquot showing the lowest RSD.
4.3.2.5 Guidelines for the interpretation of *spa* HRM data

Based on the above findings and an awareness of the possible negative impact of temperature variation across the assay plate, the following guidelines for typing isolates as same or different based on the *spa* HRM analysis on the Roche LC480 were formulated:

1. Following HRM analysis with a temperature shift of zero.
   A. Isolates should be classed as having the same *spa* type if generating difference curves which are identical or highly similarly in shape and peak which fall within an RSD <4.
   B. Isolates should be classed as having different *spa* types if generating difference curves which vary in shape and peak and have an RSD of >4.

2. In case of ambiguous results i.e.
   Isolates presenting with difference curves identical or similar in shape and peak but an RSD <4
   OR
   Isolates with difference curves identical or similar in shape, an RSD <4 but slightly differing peaks
   A. Isolates may only be typed as same or different after further analysis of results following a restoration of the automatic temperature shift function of the gene scanning software to control for temperature variation across the plate.
4.3.2.6 HRM analysis of a range of spa types

Using the above criteria for the interpretation of spa HRM data, several isolates of different spa types (Table 4.5) were then analysed. The results demonstrated the ability of the HRM technique to be used in genotyping of isolates via the spa locus (Figure 4.12). No ambiguous results occurred within these isolates which were classed by the HRM software into two groups of 2 and 3 isolates respectively and others occurring singly (singletons).
Table 4.5: List of isolates used to test the optimised spa HRM

<table>
<thead>
<tr>
<th>S/No</th>
<th>spa type</th>
<th>Isolate ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>t008</td>
<td>TS27; TS29; TS30</td>
</tr>
<tr>
<td>2.</td>
<td>t1941</td>
<td>TS1</td>
</tr>
<tr>
<td>3.</td>
<td>t128</td>
<td>NRS123; NRS248</td>
</tr>
<tr>
<td>4.</td>
<td>t386</td>
<td>TS26</td>
</tr>
<tr>
<td>5.</td>
<td>t125</td>
<td>NRS192</td>
</tr>
<tr>
<td>6.</td>
<td>t127</td>
<td>NRS229</td>
</tr>
<tr>
<td>7.</td>
<td>t021</td>
<td>NRS162</td>
</tr>
</tbody>
</table>

Figure 4.12: Normalised and zero temperature shifted difference curves of S. aureus test isolates based on the proposed criteria for HRM data analysis.

The HRM profiles were generated as described in Section 4.2.3 with a 0°C temperature shift. The figure shows accurate spa HRM genotyping with two groups of three and two isolates (indicated in black and red boxes respectively), and five singletons.

Colour code: Red (NRS192); Green (TS29); Pink (NRS248); Grey (NRS123); Yellow (TS27); Brown (TS30); Light blue (TS26); Green-Blue (NRS229), Orange (NRS162); Purple (TS1). The isolate identities are indicated in bracket.
4.3.3 Strain list for HRM Validation combining PVL and \textit{spa} types

Following the demonstration of the ability of high resolution melting to be used as a typing tool based on specific variations in \textit{lukF}, \textit{lukS} and \textit{spa} genes, the next step in the development of this technique was its validation using a larger panel of \textit{S. aureus} isolates. In order to do this, forty isolates from our collection, were assigned new identities (1 – 40) by a laboratory colleague, so that assays could be performed in a blinded manner. These isolates will subsequently be referred to as “blinded isolates”. These isolates include a range of MLST groups, \textit{spa} types and PVL status (Table 4.6) and were analysed to validate the sensitivity and specificity of this technique, as well as its potential to serve as a rapid diagnostic tool capable of determining the relatedness of PVL-positive \textit{S. aureus} isolates.
Table 4.6: Relevant characteristics of isolates used to validate the HRM assay by combining PVL and spa type data.

<table>
<thead>
<tr>
<th>S/No</th>
<th>Expected <em>lukSF-PV</em> Genotype <em>(Number of Isolates)</em></th>
<th>MLST</th>
<th>spa type</th>
<th>Isolate ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>S1FU (5)</td>
<td>ST8</td>
<td>t008</td>
<td>TS27/TS29/TS30, NRS158, TS1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST1518</td>
<td>t211</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t1941</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>SIF1 (7)</td>
<td>ST1</td>
<td>t125</td>
<td>NRS192, NRS229, NRS123/NRS248, NRS194, TS26, TS25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t127</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t128</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t175</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t386</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t3342</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST80</td>
<td>t044</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST121</td>
<td>t314</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST772</td>
<td>t345</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t657</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t310</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t852</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t3379</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t6642</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t6643</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t2554</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t690</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t6769</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>None (2)</td>
<td>ST22</td>
<td>t032</td>
<td>RSS043, RSS063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST59</td>
<td>t216</td>
<td></td>
</tr>
</tbody>
</table>

*These expected *lukSF-PV* types are based on previous associations of variations in the genes encoding the PVL toxin with specific STs as described in literature and summarised in Table 1.4.

*lukSF-PV* genotype key: S1 = G527/T663; S2 = A527/G663; F1 = A1396/A1729; F2 = G1396/A1729; FU = A1396/G1729.

S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the *lukSF-PV* locus.
4.3.4 HRM validation of the lukS gene

HRM analysis at the lukS gene resulted in 38 of the isolates classed automatically as either G₅₂⁷/T₆₆₃ or A₅₂⁷/G₆₆₃ genotypes designated S1 and S2 respectively. A subset of this data is presented in Figure 4.13. Overall, the majority of isolates (70%; 28/40), belonged to the S2 group while only 25% (10/40) belonged to the S1 group.

Following HRM analysis at the lukS locus however, two of the blinded isolates (37 and 38) could not automatically be classed as either S1 or S2. For the blinded isolate 38, while one assay duplicate was clearly identified as being ‘negative’ for the lukS gene (Figure 4.13), the second duplicate was designated as an ‘unknown’ lukS genotype. Both assay duplicates for isolate 37, were also designated as an ‘unknown’ lukS genotype. As the blank no template control in this assay was also assigned an ‘unknown’ lukS genotype by the software, the amplification and melt profiles of all three assays (blinded isolates 37, 38 and blank) were further analysed.

The melt profiles from these three assays revealed similar melt temperatures as the other blinded isolates in the test run (Figure 4.14) but significantly lower melt peaks were observed.
Figure 4.13: Representative Normalised and automatic temperature shifted difference curve profiles of the lukS locus.

The HRM profiles which were generated as described in Section 4.2.3 show a clear difference in the S1 (blue), S2 (red) and Unknown (brown) lukS profiles. S1 represents the $G_{527}/T_{663}$ genotype while S2 represents the $A_{527}/G_{663}$ genotype.
Figure 4.14: Melt curves of *lukS* locus showing lower melt peaks of negative and blank samples.

The melt curves were generated as described in Section 4.2.3. The figure shows lower melt peaks (circled in blue) of isolates 37 and 38 and assay blank no template control. Inset shows assignment of genotypes of the blinded isolates 37 and 38 (circled in red) and the blank no template control (circled in black) where brown = unknown and orange = negative.
This information is in agreement with the higher crossing point (Cp) values observed for these isolates which are represented in Figure 4.15. When compared with the Cp values (range 16.50 – 18.97) associated with the majority of the other isolates, a sub-group produced high Cp values (range 28.59 – 40.00). This sub-group included the blank no template controls and blinded isolates 37 and 38 which are PVL-negative (Figure 4.15). Cp values indicate the point at which each amplification mixture produces measurable fluorescence and is a function of the amount of DNA present at the beginning of the PCR cycling. Hence high Cp values (>30) indicate late amplification which is often non-specific – a finding that may explain discordant results in the present study. This finding highlights the need for proper monitoring and analysis of the entire amplification process and the need to take this into consideration during the interpretation of the HRM data. Proper monitoring would ensure accuracy in the downstream application as ignoring the discrepancies in melt peaks and Cp in this case would have clearly resulted in false positives.
Figure 4.15: Amplification curves showing variable crossing points generated from HRM analysis of 
lukS in a “blinded” panel of isolates.

The amplification curves were generated as described in Section 4.2.3. The figure shows late non-specific amplification 
taking place in blinded isolates 37, 38 and blank no template control (to the right of the black line) this corresponds to the 
HRM profiles with lower melt peaks as observed in Figure 4.14.
4.3.5 **HRM validation of the *lukF* gene**

For the *lukF* locus, the majority of PVL-positive isolates as expected had identical duplicates and were classed automatically as either F1 (A\textsubscript{1396}/A\textsubscript{1729} genotype), F2 (G\textsubscript{1396}/A\textsubscript{1729} genotype) or unknown. Two isolates (4 and 9) however, had inconsistent assay duplicate results. These isolates are indicated by the circles on **Figure 4.17**. In both cases one assay duplicate was assigned to a known group and the other to an unknown group. This inconsistency was however found to be eliminated by altering the automatic pre-melt normalisation temperature (circled) from a range of 71.49 – 72.49°C (**Figure 4.16 A**) to 71.19 – 72.39°C (**Figure 4.16 B**). This resulted in an equal distribution of positive isolates between both the F1 and F2 genotypes (47.5%; 19/40 of each).
Figure 4.16: Effect of automatic (A) and adjusted premelt (B) values on the consistency of HRM profiles for duplicate assay wells for isolates 4 and 9.

In A: Red circles indicate inconsistent profiles generated in duplicate wells for the same isolate before adjustment of the pre-melt normalisation temperature.

In B: Black circles indicate consistent profiles generated in duplicate wells for the same isolate after adjustment of the pre-melt normalisation temperature.

The premelt settings are indicated by blue circles on the melting curves.
4.3.6 **Summary of lukS and lukF HRM results**

Overall, the 38 PVL-positive isolates could be grouped into one of four genotypes (Table 4.7) S1F1, S1F2, S2F1 and S2F2 subsequently represented as A, B, C and D respectively. The largest number of isolates (47.3%; 18/38) belonged to the S2F2 (D) genotype which from the literature, has been associated predominantly with the ST22, ST59 and ST88 lineages (Table 1.4).

**Table 4.7: Summary of lukS and lukF HRM results of 40 blinded isolates**

<table>
<thead>
<tr>
<th>S/No</th>
<th>Detected lukSF-PV Genotype</th>
<th>Blind Identity (No of Isolates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>S1F1 (A)</td>
<td>1, 21, 22, 24, 25, 26, 31, 34, 35 (9)</td>
</tr>
<tr>
<td>2.</td>
<td>S1F2 (B)</td>
<td>28 (1)</td>
</tr>
<tr>
<td>3.</td>
<td>S2F1 (C)</td>
<td>4, 5, 6, 8, 10, 11, 13, 20, 29, 36 (10)</td>
</tr>
<tr>
<td>4.</td>
<td>S2F2 (D)</td>
<td>2, 3, 7, 9, 12, 14, 15, 16, 17, 18, 19, 23, 27, 30, 32, 33, 39, 40 (18)</td>
</tr>
</tbody>
</table>

*lukSF-PV* Genotype key: S1 = G527/T663; S2 = A527/G663; F1 = A1396/A1729; F2 = G1396/A1729; FU = A1396/G1729.

S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the *lukSF-PV* locus.
4.3.7 **HRM validation of the* spa *locus**

Based on data presented in Figure 4.17, classification of all 38 isolates using *spa* HRM was expected to generate a significant number of types, which could potentially make genotyping difficult. Therefore, further analysis of isolates was carried out by comparing difference curves within each of the PVL groupings (A, C and D) shown in Table 4.7. As group B only contained 1 isolate, this was omitted from this round of analysis.
Figure 4.17: HRM spa profiles for 20 isolates tested in triplicate.

The HRM profiles were generated as described in Section 4.2.3. The figure shows the close proximity of the HRM profiles generated in triplicates for the 20 test isolates which creates an obvious difficulty in data interpretation. Each colour represents a different isolate.
4.3.7.1 HRM spa interpretation for group A isolates

For group A (S1F1) isolates (Figure 4.18), interpretation of HRM data to differentiate isolates was relatively easy, with difference curves varying either in shape or peak maxima and clear subgroupings occurring.

Figure 4.18: HRM spa profiles generated for group A (S1F1) isolates

The HRM profiles were generated as described in Section 4.2.3. The HRM profiles show a single subgroup consisting of three isolates (22, 24 and 25) with other isolates presenting clearly distinct HRM profiles.

Colour code: Red (21); Green (22); Pink (24); Grey (25); Yellow (26); Brown (31); Light blue (34); Blue-Green (35); Blue (1). The blinded isolate numbers are shown in brackets
4.3.7.2 HRM spa interpretation for group C Isolates

For the group C isolates, two subgroups (C1 and C2) produced difference curves which were similar in shape and peak (indicated in Figure 4.19) which made differentiation of the isolates difficult or impossible. As the results were ambiguous, implementation of the guidelines described in Section 4.3.2.5 was subsequently carried out to enable successful genotyping of these subgroups.

Figure 4.19: HRM spa profiles generated for group C (S2F1) isolates.

The HRM profiles were generated as described in Section 4.2.3. The figure shows some clearly distinct HRM profiles. HRM profiles presenting similar shaped difference curves which could confound assignment of genotypes are indicated.

Colour code: Blue (1); Red (4); Green (5); Pink (6); Grey (8); Yellow (10); Brown (11); Light blue (13); Blue-Green (20); Orange (29); Purple (36). The blinded isolate numbers are shown in brackets.
For subgroup C1 (Figure 4.20), while isolates 5 and 10 were obviously identical with nearly overlapping difference curves, isolate 8 also generated a difference curve which was similar in shape and peak. Isolate 36 on the other hand, though generating a curve similar in peak, differed slightly in shape.

Despite the tendency to call isolates 5, 8 and 10 as same and isolate 36 different, a certain degree of uncertainty in the interpretation of these results still existed, hence the Section 4.3.2.5 guidelines were applied. Restoring the automatic temperature shift function at this point (as shown in Figure 4.21 B), further highlighted the relative signal difference, and showed that isolate 36 generated a >4 difference in RSD values when compared to the other three members of the cluster.

Furthermore, as isolates of identical genotype are expected to generate a <4 difference in RSD values, isolate 5 was then set as the base curve to confirm that isolates 5, 8 and 10 indeed belong to the same genotype. Without restoring the automatic temperature shift function (Figure 4.20 C), isolate 8 (coloured grey) gave a >4 RSD. Following a restoration of the automatic temperature shift function (to minimise any temperature variation across the plate), this wide difference in RSD was eliminated with all three isolates falling within a RSD of 4 (Figure 4.20 D).
Figure 4.20: HRM profiles of group C (S2F1) subgroup 1 isolates produced under varying conditions designed to clarify the genotype of closely related isolates.

The HRM profiles were generated as described in Section 4.2.3. Panel A shows the normalised and zero temperature shifted difference curves of isolates 5, 8, 10 and 36, showing HRM profiles highly similar in shape and peak. Panel B shows normalised and automatic temperature shifted difference curves of isolates 5, 8, 10 and 36 with 36 clearly differentiated with >4 difference in RSD values. Panel C and D show HRM profiles using isolate 5 as the base curve to confirm genotype of isolates 5, 8 and 10 as identical. Showing isolate 8 with a >4 difference in RSD values with zero temperature shift (Panel C) but <4 with the automatic temperature shift applied (Panel D).

Colour code: Green (5); Yellow (10), Grey (8) and Purple (36). The blinded isolate numbers are shown in brackets.
A similar interpretation procedure was carried out for the subgroup C2 of group C. This consisted of isolates 4, 6 and 13 in which isolates 4 and 6 were obviously the same, with isolate 13 being potentially different (Figure 4.21 A). In this case however, no further confirmation of relatedness could be made despite restoring the automatic temperature shift (Figure 4.21 B) and changing the base curve (Figure 4.21 D). This was as a result of the highly complex nature of the signal differences generated following the changes. Hence isolate 13 was genotyped as different (but perhaps closely related to isolates 4 and 6) based on spa HRM analysis.
Figure 4.21: HRM profiles of group C (S2F1) subgroup 2 isolates produced under varying conditions designed to clarify the genotype of closely related isolates.

The HRM profiles were generated as described in Section 4.2.3. Panel A shows normalised and zero temperature shifted difference curves of isolates 4, 6, and 13 showing profiles highly similar in shape and peak. Panel B shows the null effect on RSD value of restoring automatic temperature shift on HRM profile of isolates 4, 6, and 13. Panels C and D show HRM profiles using isolate 4 as the base curve to confirm genotype of isolates 4, 6 and 13. The data shows isolate 13 with a >4 difference in RSD value with zero temperature shift (Panel C) which could not be resolved due to the amorphous curves generated with the automatic temperature shift applied (Panel D).

Colour code: Red (4); Pink (6) and Blue (13). The blinded isolate numbers are shown in brackets.
4.3.7.3 HRM spa interpretation for group D isolates

With the exception of one subgroup, subgroup D3 (Figure 4.22), genotyping of group D isolates proved relatively straightforward with clear demarcations occurring between profiles generated by the different clusters of isolates (Figure 4.22).

![Figure 4.22: HRM spa profiles generated for group D (S2F2) isolates.](image)

The HRM profiles were generated as described in Section 4.2.3. The figure shows some clearly distinct HRM profiles. Circled are the HRM profiles for subgroup D3 which required additional interpretation.

Colour code: Red (2); Green (3); Pink (7); Grey (18); Yellow (12); Brown (40); Light blue (14); Blue-Green (9); Orange (17); Purple (32); Green 2 (19); Pink 2 (27); Grey 2 (32); Yellow 2 (39); Brown 2 (23); Light blue 2 (15); Blue-Green 2 (16); Orange 2 (30); Purple (36). The blinded isolate numbers are shown in brackets.
Further analysis of subgroup D3 HRM profiles (Figure 4.23) provided some clarification with respect to their genotyping. An expanded view of the subgroup highlighted variations in the HRM profiles with differences in both shape and peak maxima apparent.

![Normalized and Temp-Shifted Difference Plot](image)

**Figure 4.23: HRM spa profiles of group D (S2F2) subgroup 3 only.**

The HRM profiles were generated as described in Section 4.2.3. The figure shows a more detailed analysis of HRM profiles of subgroup D3 further highlighting differences in shape and peak of the difference curves to enable further genotyping.

Colour code: Green (3); Pink (7); Grey (18); Light blue (14); Green 2 (19); Pink 2 (27); Light blue 2 (15); Blue-Green 2 (16); Orange 2 (30). The blinded isolate numbers are shown in brackets.

While the remaining isolates in subgroup D3 (3, 7, 18, 19 and 27) produced obviously varied difference curves (Figure 4.24), the <4 difference in RSD values provided a strong indication of identical *spa* genotypes. Hence, the guidelines prescribed in Section 4.3.2.5 were applied. Restoring the automatic temperature shift function of the software at this point however, simply resulted in curves which were more complex and difficult to interpret (not shown). No further confirmation could therefore be carried out, hence the isolates were tentatively classed as same.
Figure 4.24: Subgroup D3 (S2F2) isolates showing complex HRM difference curves

The HRM profiles were generated as described in Section 4.2.3. Figure shows complex HRM curves which could not be visually assigned to unique groups.

Colour code: Green (3); Pink (7); Grey (18); Green 2 (19); Pink 2 (27); Orange 2 (30). The blinded isolate numbers are shown in brackets.
4.3.8 Compilation of HRM results

Taking all of the above into consideration, the spa HRM resulted in 7 different spa types identified for PVL group A (S1F1). This was comprised of 1 subgroup of 3 isolates and 6 singletons. PVL group B (S1F2) was composed of only a single isolate, hence 1 spa type. For PVL group C (S2F2), 7 spa types were identified. These were comprised of 2 subgroups of 2 and 3 isolates each and 5 singletons. PVL group D (S2F2) had 11 spa types. These were comprised of 3 subgroups of 2, 3 and 5 isolates respectively and 8 singletons (Table 4.8). Overall, employing HRM analysis as a rapid genotyping tool for the 38 PVL-positive isolates resulted in a total of 26 distinct genotypes (Table 4.9). The total number of genotypes detected by HRM analysis of both lukSF-PV and spa loci in the present study however, differed from the expected number of genotypes. Based on known spa types and STs, a total of 23 distinct genotypes were expected. These were expected to be contained within four PVL groups (S1F1, S2F1, S2F2 and S1FU) comprising 6, 5, 9 and 3 genotypes respectively.
Table 4.8: Compilation of *lukSF-PV* and *spa* HRM results to create unique code for each genotype

<table>
<thead>
<tr>
<th><em>lukSF-PV</em> Genotype</th>
<th><em>spa</em> HRM groupings</th>
<th>Corresponding Blinded Isolate identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1F1 (A)</td>
<td>A1</td>
<td>22/24/25</td>
</tr>
<tr>
<td></td>
<td>A2-7 (Singletons ***)</td>
<td>1/26/31/34/35/21</td>
</tr>
<tr>
<td>S1F2 (B)</td>
<td>B1</td>
<td>28</td>
</tr>
<tr>
<td>S2F1 (C)</td>
<td>C1</td>
<td>4/6</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>5/8/10</td>
</tr>
<tr>
<td></td>
<td>C3-7 (Singletons)</td>
<td>11/13/20/29/36</td>
</tr>
<tr>
<td>S2F2 (D)</td>
<td>D1</td>
<td>14/15</td>
</tr>
<tr>
<td></td>
<td>D2</td>
<td>12/39/40/</td>
</tr>
<tr>
<td></td>
<td>D3</td>
<td>3/7/18/19/27</td>
</tr>
<tr>
<td></td>
<td>D4-11 (Singletons)</td>
<td>2/9/16/17/23/30/32/33</td>
</tr>
</tbody>
</table>

*The *lukSF-PV* genotype is as detected in the present study and described in Table 4.7*

**The *spa* groups summarise the results generated in Section 4.3.7, with subgroups within each *lukSF-PV* genotype numbered.**

***Singletons indicate groups composed of only a single isolate***

*lukSF-PV* genotype key: S1 = G\text{527}/T\text{663}; S2 = A\text{527}/G\text{663}; F1 = A\text{1396}/A\text{1729}; F2 = G\text{1396}/A\text{1729}; FU = A\text{1396}/G\text{1729}.

S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the *lukSF-PV* locus.
Table 4.9: Comparison of experimentally generated HRM genotypes with predicted results

<table>
<thead>
<tr>
<th>lukSF-PV genotype</th>
<th>Expected* Number of Isolates (No of spa types)</th>
<th>Actual** Number of Isolates (No of spa types)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1F1</td>
<td>7 (6)</td>
<td>9 (7)</td>
</tr>
<tr>
<td>S1F2</td>
<td>0 (0)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>S2F1</td>
<td>12 (5)</td>
<td>10 (7)</td>
</tr>
<tr>
<td>S2F2</td>
<td>14 (9)</td>
<td>18 (11)</td>
</tr>
<tr>
<td>S1FU</td>
<td>5 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total</td>
<td>38 (23)</td>
<td>38 (26)</td>
</tr>
</tbody>
</table>

*The expected types are predicted based on previous associations of variations in the genes encoding the PVL toxin with specific STs as described in published literature and summarised in Table 1.4.

**The actual predicted types are as generated in the present study.

lukSF-PV genotype key: S1 = G_{527}/T_{663}; S2 = A_{527}/G_{663}; F1 = A_{1396}/A_{1729}; F2 = G_{1396}/A_{1729}; FU = A_{1396}/G_{1729}.

S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the lukSF-PV locus.
These results indicate that using HRM as a typing tool resulted in 100% typeability of the PVL-positive *S. aureus* isolates. HRM typing yielded a similar number of types and distribution within types (26 types in all) when compared with the combination of MLST and *spa* types (23 types in all). Furthermore, comparing the discriminatory power of both techniques as previously described (Hunter and Gaston, 1988) using Formula 1 revealed a comparable index of discrimination between both typing methods i.e. HRM and the MLST/*spa* typing. The indices of discrimination (D) were 0.947 and 0.97 for the MLST/*spa* combination and HRM typing respectively.

\[
D = 1 - \frac{1}{N(N - 1)} \sum_{j=1}^{S} n_j(n_j - 1)
\]

**Formula 1: Parameters used to ascertain discriminatory power.**

(Where D is the index of discrimination, N is the population size; S is number of types and n is the distribution of strains within types)

These figures were generated by applying the information from both typing schemes to the formula. In both cases the population size (N) was identical (38).

With regards to the MLST/*spa* combination typing with number of types (s) equal to 23 and strain distribution (n) varying from 7 to 1 (\(n_1 = 7, n_2 = 5, n_3 = 3, n_4 = 2\) and \(n_8\) to \(n_{25} = 1\); D could therefore be calculated thus:

\[
1 - [(7 \times 6 + 5 \times 4 + 3 \times 2 + 2 \times 1 + 2 \times 1 + 1 \times 1 + 1 \times 0 \ldots + 1 \times 0)/(38 \times 37)] =
1 - (74/1406) = 0.947.
\]
Similarly for the HRM data with $s = 26$ and $n$ values ranging from 5 to 1 ($n_1 = 5$, $n_2$ to $n_4 = 3$, $n_5$ to $n_6 = 2$ and $n_7$ to $n_{26} = 1$) the D could therefore be calculated thus:

$$1 - \frac{(5 \times 4 + 3 \times 2 + 3 \times 2 + 2 \times 1 + 2 \times 1 + 1 \times 0 \ldots + 1 \times 0)}{(38 \times 37)} = 1 - \frac{42}{1406} \approx 0.970.$$
4.3.8.1 **Data Correlation**

With each blind isolate assigned a specific code based on a combination of *lukS*, *lukF*, and *spa* HRM data (represented by a letter and a number as described in Table 4.8), a further step was taken to validate the ability of HRM to be used as a rapid diagnostic tool. This involved matching up the blinded isolates to the original isolate list. Prior to this process however, sequencing of both *lukS* and *lukF* genes of all 38 PVL-positive isolates was carried out. The sequencing was done to confirm the current PVL data which had so far been inferred from the sequence type of isolates, based on published literature (Table 1.4). The sequences obtained from most isolates correlated with the expected data but some exceptions occurred. Of the 38 PVL-positive isolates, 9 showed allelic variations in the *lukSF-PV* locus which differed from the expected genotype (Table 4.10).

Based on the original and updated information (Table 4.6 and Table 4.10), the final step in this process involved predicting the actual isolate identity of each blinded isolate based on the experimental data obtained. This information is presented in Table 4.11.
Table 4.10: Variations from expected PVL genotypes detected via sequencing

<table>
<thead>
<tr>
<th>S/No</th>
<th>ST (expected genotype&lt;sup&gt;1&lt;/sup&gt;)</th>
<th>Isolate identity</th>
<th>Actual <em>lukSF-PV</em> genotype&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>ST1 (S1F1)</td>
<td>TS25, TS26, NRS229, NRS194</td>
<td>S2F1, S1FU, SIFU, S2F2</td>
</tr>
<tr>
<td>2.</td>
<td>ST8 (S1FU)</td>
<td>NRS158</td>
<td>S2F2</td>
</tr>
<tr>
<td>3.</td>
<td>ST30 (S2F1)</td>
<td>TS12, NRS162, RSS289, RSS290</td>
<td>S2F2, S1F2, S2F2, S2F2</td>
</tr>
</tbody>
</table>

<sup>1</sup>The expected genotypes are based on previous associations of variations in the genes encoding the PVL toxin with specific STs as described in literature and summarised in Table 1.4.

<sup>2</sup>The actual PVL genotype was determined by sequencing of the *lukSF-PV* locus

*lukSF-PV* genotype key: S1 = G<sub>527</sub>/T<sub>663</sub>; S2 = A<sub>527</sub>/G<sub>663</sub>; F1 = A<sub>1396</sub>/A<sub>1729</sub>; F2 = G<sub>1396</sub>/A<sub>1729</sub>; FU = A<sub>1396</sub>/G<sub>1729</sub>.

S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the *lukSF-PV* locus.
Table 4.11: Linking of blinded isolate identity to experimentally determined isolate identities

<table>
<thead>
<tr>
<th>lukSF-PV Genotype</th>
<th>spa HRM groupings</th>
<th>Blind Identity</th>
<th>Proposed Isolate Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1F1 (A)</td>
<td>A1</td>
<td>22/24/25</td>
<td>TS27/TS29/TS30</td>
</tr>
<tr>
<td></td>
<td>A2-7 (sgls)</td>
<td>1/26/31/34/35/21</td>
<td>TS1/TS26/NRS123/NRS192/NRS229/NRS248</td>
</tr>
<tr>
<td>S1F2 (B)</td>
<td>B1</td>
<td>28</td>
<td>NRS162</td>
</tr>
<tr>
<td>S2F1 (C)</td>
<td>C1</td>
<td>4/6</td>
<td>TS2/TS5</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>5/8/10</td>
<td>TS7/TS15/TS16/</td>
</tr>
<tr>
<td></td>
<td>C3-7 (sgls)</td>
<td>11/13/20/29/36</td>
<td>TS8/TS25/NRS185/NRS255</td>
</tr>
<tr>
<td>S2F2 (D)</td>
<td>D1</td>
<td>14/15</td>
<td>TS6/TS9</td>
</tr>
<tr>
<td></td>
<td>D2</td>
<td>12/39/40/</td>
<td>TS12/RSS289/RSS290</td>
</tr>
<tr>
<td></td>
<td>D3</td>
<td>3/7/18/19/27</td>
<td>TS13/T14/TS18/TS20/TS24</td>
</tr>
</tbody>
</table>

*The proposed isolate identity shown in this Table is based on the information provided in Table 4.6 and the lukSF-PV sequencing data in Table 4.10. sgls: singletons

lukSF-PV genotype key: S1 = G_{527}/T_{663}; S2 = A_{527}/G_{663}; F1 = A_{1396}/A_{1729}; F2 = G_{1396}/A_{1729}; FU = A_{1396}/G_{1729}.
S1 and S2 represent allelic variations at positions 527 and 663 while F1, F2 and FU represent allelic variations at positions 1396 and 1729 of the lukSF-PV locus.
Combining the expected and predicted results at the *lukSF-PV* locus, the assay appears to be 100% sensitive and 81.58% specific. The assay was able to correctly identify all 38 isolates as PVL-positive. However, the 6 isolates belonging to the S1FU group (based on data from Tables 4.7 and 4.11) were misclassified as belonging to group S1F1, while a single S2F1 isolate (TS17) was misclassified as S2F2.

While combining the data from the analysis of the *spa* locus did improve the discriminative ability of this HRM based typing method and enabled further delineation within groups, the actual positive predictive value of this technique was however poor. The predictive value was ascertained by the ability to accurately link a blind identity to a known isolate. This low predictive value was highlighted with the six *spa* groups containing 2 or more isolates (Table 4.12). An accurate prediction of the actual strain identity of the blinded isolate was possible in only 2 of these 6 groups (33.3%).
Table 4.12: *spa* HRM predictive ability

<table>
<thead>
<tr>
<th>S/No</th>
<th>Blind Identity groupings</th>
<th>Proposed Isolate Identity</th>
<th>Actual Isolate Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>22/24/25</td>
<td>TS27/29/30</td>
<td>TS27/29/30</td>
</tr>
<tr>
<td>2.</td>
<td>4/6</td>
<td>TS2/TS5</td>
<td>TS15/TS16</td>
</tr>
<tr>
<td>3.</td>
<td>5/8/10</td>
<td>TS7/TS15/TS16</td>
<td>TS2/TS5/TS8</td>
</tr>
<tr>
<td>4.</td>
<td>14/15</td>
<td>TS6/TS9</td>
<td>TS14/TS20</td>
</tr>
<tr>
<td>5.</td>
<td>12/39/40</td>
<td>TS12/RSS289/RSS290</td>
<td>TS12/RSS289/RSS290</td>
</tr>
<tr>
<td>6.</td>
<td>3/7/18/19/27</td>
<td>TS13/TS14/TS18/TS20/TS24</td>
<td>TS13/TS18/TS24</td>
</tr>
</tbody>
</table>

*The groupings for blinded isolates were ascertained experimentally from results of HRM analysis at both *lukSF-PV* and *spa* locus.*

**The proposed isolate identity indicates the predicted identity of each blinded isolate based on results of the HRM analysis**

***The actual isolate identity indicates the true strain identity of each isolate on the blinded list (Appendix 3)*

The two groups of isolates whose genotypes were accurately predicted are depicted in red.
In three of the seven expected spa groups (Table 4.13), all members of the group were correctly identified as belonging to the same group. For the TS2/TS5 (t657) group though, typing via HRM resulted in an isolate with an unrelated spa type (t345, TS8) misidentified as part of this group. A partially correct identification of group members belonging to the same spa type was however made in two cases. In the case of the TS7/TS15/TS16 group, a single member (TS7) was misidentified as being different. Isolates that were mistakenly assigned to the different spa groups belonged to the group C clusters whose HRM profiles had initially provided conflicting results that required further analysis (Figure 4.21 and Figure 4.22).

These findings highlight the limitations of the use of the spa HRM. While the issue of temperature variability across the assay plate did not negatively impact on the discriminatory ability of the technique as a whole, the subjectivity introduced into the data interpretation due to temperature variation poses a significant problem in the predictive accuracy of this technique using the spa locus.
Table 4.13: Accuracy of *spa* HRM groupings compared to known *spa* types

<table>
<thead>
<tr>
<th>S/No</th>
<th>Known <em>spa</em> groups</th>
<th>Correctly grouped as same</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS27/29/30</td>
<td>YES</td>
</tr>
<tr>
<td>2.</td>
<td>TS12/RSS289/290</td>
<td>YES</td>
</tr>
<tr>
<td>3.</td>
<td>TS2/TS5</td>
<td>YES</td>
</tr>
<tr>
<td>4.</td>
<td>TS7/15/16</td>
<td>PARTIALLY</td>
</tr>
<tr>
<td>5.</td>
<td>TS13/TS14/TS18/TS20/TS24</td>
<td>PARTIALLY</td>
</tr>
<tr>
<td>6.</td>
<td>NRS123/248</td>
<td>NO</td>
</tr>
<tr>
<td>7.</td>
<td>TS6/TS9</td>
<td>NO</td>
</tr>
</tbody>
</table>

*Response key: ‘YES’ = All members of the group identified as same; ‘PARTIALLY’ = Some members identified as same and some misidentified as different; ‘NO’ = All members misidentified as different*
4.3.9 **Effect of mutation position on outcome of HRM analysis**

Following the ability of the HRM analysis to differentiate the two *lukS* genotypes possessing two SNPs with potentially identical melt temperatures (a G – A SNP at position 527 and a T – G SNP at position 663), the present study went on to further explore the contribution of SNP position in differentiating between these potentially neutralising nucleotide changes in the *lukS* locus.

To do this, four custom DNA sequences were synthesized (produced by DNA2.0) having an identical base pair composition overall but a G/T to A/G allelic variations at varying positions within the sequence (Table 4.14). These were subjected to HRM analysis using the original sequence as a control. The gene scanning software was able to again place the original *lukS* G\(_{527}/T_{663}\) allelic variant and the G\(_{527}/T_{663}\) variant into two separate genotypes (Figure 4.25), thus affirming the high sensitivity and reproducibility of the initial HRM analysis. Interestingly however, all A/G variants (which have exactly identical base pair composition) were not classed together, suggesting a position dependent effect of allele variation on HRM sensitivity.

The four A/G allelic variants were classed into three different groups by the gene scanning software. The *lukS* variant A was classed as identical with the original *lukS* G\(_{527}/T_{663}\) sequence. While the *lukS* variant C and the original *lukS* A\(_{527}/G_{663}\) variant were placed in the same genotype, the *lukS* variant B was grouped entirely on its own. A visual inspection of the normalised and automatic temperature shifted difference curve of this variant however shows a close relatedness to the *lukS* variant A group as noted by a similar shaped difference curve (Figure 4.25).
Table 4.14: Single nucleotide polymorphisms in the 4 different test lukS sequences in comparison to a parent sequence

| lukS Variant A | GTGGTCCATCAACAGGAGGT | AATGATTCATTTTAATTATTCAAAAAACAATTTAGTTA | TAATCAACAAAAACTATATCGTGAATGAGCTCAAAATTCAAAAAAGTTTCAA | TGGGAATAAAAAAGCTAATTCTATATCAACATTAGTTAAATATGTCGGACATG | ATCCAAAATTTTTGTTGATATAAAACCATTATATGCTAAAAATCCGAGAGACTATTTTGTTCCAGACAAATGAATTACCCCACTATAGTACACAGTGGTTTCAATCCTTCAGACACAGTGGTTTCAATCCTTC
| lukS Variant B | GTGGTCCATCAACAGGAGGT | AATGATTCATTTTAATTATTCAAAAAACAATTTAGTTA | TAATCAACAAAAACTATATCGTGAATGAGCTCAAAATTCAAAAAAGTTTCAA | TGGGAATAAAAAAGCTAATTCTATATCAACATTAGTTAAATATGTCGGACATG | ATCCAAAATTTTTGTTGATATAAAACCATTATATGCTAAAAATCCGAGAGACTATTTTGTTCCAGACAAATGAATTACCCCACTATAGTACACAGTGGTTTCAATCCTTCAGACACAGTGGTTTCAATCCTTC
| lukS Variant C | GTGGTCCATCAACAGGAGGT | AATGATTCATTTTAATTATTCAAAAAACAATTTAGTTA | TAATCAACAAAAACTATATCGTGAATGAGCTCAAAATTCAAAAAAGTTTCAA | TGGGAATAAAAAAGCTAATTCTATATCAACATTAGTTAAATATGTCGGACATG | ATCCAAAATTTTTGTTGATATAAAACCATTATATGCTAAAAATCCGAGAGACTATTTTGTTCCAGACAAATGAATTACCCCACTATAGTACACAGTGGTTTCAATCCTTCAGACACAGTGGTTTCAATCCTTC

Original lukS A\textsubscript{527}/G\textsubscript{663} Variant**

GTGGTCCATCAACAGGAGGT | AATGATTCATTTTAATTATTCAAAAAACAATTTAGTTA | TAATCAACAAAAACTATATCGTGAATGAGCTCAAAATTCAAAAAAGTTTCAA | TGGGAATAAAAAAGCTAATTCTATATCAACATTAGTTAAATATGTCGGACATG | ATCCAAAATTTTTGTTGATATAAAACCATTATATGCTAAAAATCCGAGAGACTATTTTGTTCCAGACAAATGAATTACCCCACTATAGTACACAGTGGTTTCAATCCTTCAGACACAGTGGTTTCAATCCT

The yellow highlighting indicates the location of the primer binding sites for HRM analysis.
The SNP\textsubscript{A} positions in the sequences are highlighted in blue and SNP\textsubscript{B} in red.

*Sequence source: Strain NN1, GenBank accession no. AB186917
**Sequence source: Strain MW2, GenBank accession no. BA000033

210
Figure 4.25: Normalised and automatic temperature shifted difference curves of *lukS* locus with identical DNA composition but varying SNP location

The HRM profiles were generated as described in Section 4.2.3. Panel A shows the automatic classification of samples (each colour represents a unique group), maintaining the ability of the technique to differentiate between the original A\textsubscript{527}/G\textsubscript{663} allele from the G\textsubscript{527}/T\textsubscript{663} allele (arrowed). Panel B shows the samples as originally colour coded prior to the automatic classification indicating the sample identities.

Colour code for Panel B: Blue = Original *lukS* A\textsubscript{527}/G\textsubscript{663} variant; Yellow = *lukS* Variant C; Pink = *lukS* Variant B; Red = *lukS* Variant A; Orange = Original *lukS* G\textsubscript{527}/T\textsubscript{663} sequence.

No identifiable pattern appears to account for the grouping/genotyping of these five sequences, as HRM products with the most similar melt temperatures were not necessarily placed in the same group (Table 4.15).
Table 4.15: Variable characteristics of the *lukS* test sequences

<table>
<thead>
<tr>
<th>Sequence Types</th>
<th>SNP_A Proximal* to forward primer</th>
<th>SNP_A GC rich** neighbours</th>
<th>SNP_B Proximal to Reverse primer</th>
<th>SNP_B GC rich neighbours</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original <em>lukS</em> G\textsubscript{527}/T\textsubscript{663} sequence</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>75.99</td>
</tr>
<tr>
<td>Original <em>lukS</em> A\textsubscript{527}/G\textsubscript{663} variant</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>76.03</td>
</tr>
<tr>
<td><em>lukS</em> Variant A</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>75.79</td>
</tr>
<tr>
<td><em>lukS</em> Variant B</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>75.73</td>
</tr>
<tr>
<td><em>lukS</em> Variant C</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>75.77</td>
</tr>
</tbody>
</table>

*Proximal is defined as a mutation occurring within 16 bp of the primer

**GC rich neighbours describe mutations of which 3 of the 4 immediate adjacent base pairs are either a G or a C

NA: Not Applicable

Mutation A: G to A mutation

Mutation B: T to G mutation

Tm (°C): Melt temperature of each sequence following HRM
4.4 Discussion

HRM has been shown to exhibit a remarkably high level of sensitivity when compared to traditional melt curve analysis (Wittwer et al., 2003). This is thought to depend in part on improved dye chemistry. The saturating dyes used in HRM are neither toxic nor inhibitory at high concentrations and are therefore able to bind to saturation. This prevents dye relocation during the melting phase, hence leading to improved sensitivity (Wittwer et al., 2003, Erali et al., 2008). The present study therefore set out to explore the potential of HRM as a rapid diagnostic tool in the characterization of PVL-positive S. aureus isolates based on the premise that this technique would be able to accurately differentiate between the double mutations contained in the lukS and lukF gene fragments.

In the present study, the totally different HRM curves generated for PCR amplicons from four standard NARSA isolates, which theoretically possess identical melt temperatures (Figure 4.2 A), confirmed the increased discriminatory power of this technique compared to that obtained using a traditional melt curve analysis. The HRM melt curves of PCR amplicons from the lukF locus however differed significantly from those of lukS. The lukF products specifically exhibited a complex melt profile composed of two melt domains. This phenomenon has previously been reported in amplicons >400 bp in length (Wittwer et al., 2003) and results from the melting of AT-rich regions at lower temperatures than the GC-rich regions.

Improved dye chemistry alone is however insufficient to generate the additional information provided by the HRM technique, as observed by analysis of the raw HRM data (Figure 4.2 A). This data was only able to provide useful information after
further analysis using the Roche Gene Scanning Software Version 1.5.0 (Figure 4.1 D). Most HRM instruments are designed to work with specific software which controls up to three functions – normalisation, temperature shift and difference curve generation. The normalisation function which corrects for varying template concentration can be described as a shifting on the Y – axis and enables comparison of the different melt curves, thereby highlighting subtle differences. The second function – temperature shift – corrects for slight temperature variations across the plate (Reja et al., Wittwer et al., 2003). Difference curves may then be generated by subtracting the normalised and temperature shifted curves from that of a reference ‘base curve’, a process which makes it easier to visualise small variations between amplicons.

Data generated by HRM analysis, was able to accurately distinguish the two allelic variants in the lukS locus which potentially balanced each other out thermodynamically (G<sub>527</sub> – A<sub>527</sub>, T<sub>663</sub> – G<sub>663</sub>). These results were obtained without any modifications to the automatic software settings. This finding is contrary to a previous report in which an amplicon with a possibly neutralising double mutation could only be detected following changes in instrument settings for normalization, temperature shift and sensitivity (Pietzka et al., 2009). Another study involving a C – T, A – C double mutation in an 88 bp fragment of the gyrA gene of Bacillus anthracis may suggest a possible variable effect of study sequence, as these mutations could not be detected using HRM analysis (Loveless et al., 2010).

In the present study however, an extra T – A mutation at nucleotide 470 of lukS detected by DNA sequence analysis in three of the ST30 isolates (data not shown) could not be detected by HRM. Several factors have been shown to affect both the
sensitivity and specificity of HRM. Most notable are the amplicon length, GC content and class of mutation i.e. specific base change (Reed and Wittwer, 2004, Liew et al., 2004, Graham et al., 2005, Palais et al., 2005, Krypuy et al., 2006). T/A mutations which belong to the previously described class 4 SNPs, are generally known as one of the more difficult SNP classes to detect, with a T_m shift of <0.1°C. While occurrence of this SNP has been shown to be readily detected in an amplicon of <50 bp (Liew et al., 2004), generally, small variations in T_m are less easily detected with increasing amplicon lengths. The relatively large size of the lukS amplicon (272 bp) may therefore explain the failure to detect the additional T/A mutation in the ST30 isolates. This effect of amplicon length on the sensitivity of HRM in detection of mutations is applicable across all SNP classes. Hence, amplicon lengths of less than 300 bp are currently recommended to ensure highest sensitivities (Krypuy et al., 2006, Reed and Wittwer, 2004).

A further variable which could influence discriminatory power of HRM – mutation position within the amplicon – was also investigated using the lukS locus. One previous study carried out in 2004 had investigated the effect of mutation position on HRM sensitivity (Reed and Wittwer, 2004). However, the approach in this Reed and Wittwer study involved altering the mutation position by changing primer binding sites to create different amplicons of similar length. This process would however invariably create products differing in GC content (however slight) which could influence HRM profiles. The present study analysed lukS sequences which were identical in length and GC content but differed singly in the mutation position. The results showed that the position of the mutation did indeed affect discriminatory power of HRM for the lukS locus (Section 4.3.9). While the sequences tested,
theoretically have the same basic melt temperatures \( (T_m) \), due to their identical base composition and hence GC content, based on the nearest neighbour model (von Ahsen et al., 1999, Peyret et al., 1999) the \( T_m \) might differ. This model theorises that the melt behaviour of a DNA amplicon is dependent not just on its length and GC content but also on arrangement. The identity and orientation of neighbouring base pairs could affect the thermal stability of a given base pair and the nearest neighbour model has been found to give a more accurate \( T_m \) prediction (SantaLucia, 1998, SantaLucia et al., 1996, von Ahsen et al., 1999).

In the present study, preliminary results based on mutations in the \( lukSF-PV \) locus have established HRM as a potentially useful technique in the rapid identification and genotyping of PVL-positive \( S. aureus \) isolates. Results however showed that the level of discrimination provided using this locus, was limited. Only three genotypes were represented within the initial test population as opposed to the six STs defined by MLST. As this factor could have a significant negative impact on the value of this technique, this study then went further to explore the possibility of increasing the discriminatory power by incorporating a second locus into the assay.

The \( spa \) gene with its highly variable nature and current successful application in the typing of \( S. aureus \) isolates seemed a logical choice as the second locus to improve the discrimination of this present study. Most published studies though, have only used HRM to detect slight variants such as SNPs, with a few studies analysing short insertions or deletions in a product of specific length (Ciammaruconi et al., 2009, Parant et al., 2009). Applying HRM to the \( spa \) locus would however involve the differentiation of products of variable lengths, due to the nature of the \( spa \) gene. This
application of HRM to length polymorphisms has not been widely employed (Parant et al., 2009, Chan et al., 2009, Stephens et al., 2008).

One previous study (Stephens et al., 2008) had explored the use of HRM in detecting different *spa* types using the Corbett Rotor-Gene 6000 HRM platform, but no further reports of the potential application of this method on other HRM platforms have been published. In the present study using HRM analysis at the *spa* locus on a Roche LC480, an initial hurdle was encountered right at the start of data analysis. Although temperature shift of the data is often important to eliminate slight temperature differences between samples, with respect to the *spa* gene, the use of this function only served to complicate result interpretation. Such a phenomenon was neither reported by Stephens et al. (2008) or in a second similar study involving the detection of variation in a clustered, regularly interspaced short-palindromic-repeat (CRISPR) (Price et al., 2007). CRISPR is structurally similar to the *spa* gene in that it is composed of variable numbers of repeats. Both of these published studies made use of the Corbett Rotor-Gene 6000. This differs from the Roche LC480 used in the present study in that it lacks the temperature shift function. These variations therefore possibly represent differences in the HRM platform.

That the choice of HRM platform in use has an effect on result outcome has been previously observed (De Leeneer et al., 2008). As would be expected considering the variations in make and design, the various HRM platforms differ in dye chemistry, software used, mode of data acquisition, sample vessel (capillaries verse plates) and mode of heat distribution. These differences in turn potentially impact on result output. This was highlighted by the ability of HRM to distinguish between two highly
similar SNPs (c.3113A>G and c.3119G>A) in a 378 bp BRCA1 gene fragment in a 2009 study (van der Stoep et al., 2009). These findings were in contrast to a 2008 study in which these two SNPs were classed in the same group (De Leeneer et al., 2008). These studies simply differed in the HRM platform used. While van der Stoep et al. (2009) used a 96-well LightScanner (Idaho Technology), De Leeneer et al. (2008) made use of LightCycler 480 (Roche). Furthermore, De Leeneer and colleagues comparing the use of two different HRM platforms in screening for BRCA1 and BRCA2 mutations reported a higher degree of specificity and sensitivity using the 96-well LightScanner as opposed to the Roche LC480 (De Leeneer et al., 2008).

One of the major factors impacting on scanning sensitivity, specificity and accuracy of genotyping, is the degree of temperature uniformity exhibited by these instruments. Two previous studies (Herrmann et al., 2006, Herrmann et al., 2007) had set out to compare the use of HRM for mutation scanning and genotyping on the six different HRM platforms then available (Figure 4.26). These studies noted a wider ‘across plate’ temperature variation in instruments using a heat block system as opposed to air based systems.
Figure 4.26: Illustration of across plate temperature variation observed in the different HRM platforms

Reproduced from: Herrmann et al. (2007). Panel C on the figure shows the dynamic melt profile of the different HRM instruments. The study was carried out using variants of a 110 bp fragment of the β-globin gene.
The Roche LC480 utilized in this present study is a heat block system and might be expected to exhibit well to well temperature variation across the plate. This problem was clearly highlighted in the present study with the position dependent variation in amplification efficiency (Figure 4.8). This phenomenon impacted negatively on interpretation of the spa HRM data (Figure 4.9), by assigning identical aliquots analysed at varying well positions as different. This finding thereby created a need to take account of sample position in interpretation of the spa HRM data, hence limiting the use of this technique at the spa locus.

As there was no indication that well to well temperature variation negatively impacted on interpretation of the lukSF-PV HRM data, the spa data highlights the importance of choosing the appropriate locus for analysis by a particular typing system. The lukSF-PV locus potentially has only a few variants. In the present study, these could therefore be identified using strains of known genotypes as internal standards. In contrast, over 12,000 spa types currently exist (http://spa.ridom.de/spatypes.shtml, assessed 2nd August 2013), making the use of internal standards unrealistic, creating a need for strain to strain comparison and thus enabling the well to well variation to have a negative impact on typing outcome.

However, variation across the heating block may not be the only factor responsible for the issues raised in spa HRM data interpretation. The study by Price and colleagues using the Corbett Rotor-Gene 6000 identified an unexpected finding in relation to the dye used in the assay. Surprisingly, better results were obtained using the non-saturating SYBR green I dye as opposed to the saturating SYTO9 dye. The SYTO9
dye resulted in less reproducible results which resulted in several duplicates being classed as different (Price et al., 2007).

Overall the results presented in this chapter highlight the potential of HRM as a rapid typing alternative to the sequence based techniques currently in use. This potential was however limited by the limited discrimination using the lukSF-PV locus and the subjectivity needed for the interpretation of data for the spa locus. While HRM analysis of both loci characterised in the present study allowed detection of differences between various genotypes, overall, the poor predictive ability raises queries on its suitability as a typing tool based on the lukSF-PV and spa loci. The arising issues, rather than being only a function of the technique itself, appear to be a sum of the technique plus the HRM platform in use as well as the reagents. This present study therefore provides an indication of the potential value of HRM as a rapid diagnostic tool but also warns of possible hurdles in fulfilling this potential and the importance of selecting appropriate loci for analysis.
Chapter Five

5 Development and Validation of an Enzyme-Linked ImmunoSorbent Assay for the detection and quantification of PVL and HLA toxins

5.1 Introduction

Currently, diagnosis of infections caused by PVL-positive strains of \textit{S. aureus}, as documented in guidelines provided by the HPA (HPA, 2008), is based on clinical suspicion (recurrent boils, abscesses, necrotising skin and soft tissue infection, community-acquired necrotising pneumonia) and an antibiotic profile of gentamicin/trimethoprim resistance previously associated locally with the presence of \textit{lukSF-PV} (Boakes et al., 2011b). Actual confirmation that these cases are caused by PVL-positive \textit{S. aureus} however, is usually dependent on toxin gene profiling carried out by the HPA Staphylococcus Reference Unit (SRU). These tests are carried out daily and expected to be completed within a working day. They however only provide information regarding the presence or absence of the encoding genes, rather than the level of toxin gene expression. Information on toxin gene expression could act as a pointer to potential severity and prognosis of disease, enabling a more informed treatment strategy, thereby impacting on patient care.

Presently, laboratory identification of \textit{S. aureus} involves a range of methods. Following culture, isolates may be identified as \textit{S. aureus} using various phenotypic methods such as the production of staphylococcal protein A and heat-stable nuclease.
Tube and/or slide coagulase tests then commonly serve as confirmatory tests (Fonsale et al., 2004, Brown et al., 2005). In addition, rapid commercial biochemical test systems have been developed which detect both clumping factor and staphylococcal protein A. Some tests additionally detect S. aureus group specific cell surface antigens (bioMerieux Slidex Staph Plus and Murex Diagnostics Ltd Staphaurex Plus) or the S. aureus capsular polysaccharide (Sanofi Diagnostics Pastorex Staph-Plus) (van Griethuysen et al., 2001, Smole et al., 1998).

More expensive molecular test systems for the identification of S. aureus also exist. These have commonly been designed to target species-specific genes, some of which include genes encoding the nuclease, coagulase and protein A, as well as femA, femB and 16S rRNA (Brown et al., 2005, Brakstad et al., 1992, Martineau et al., 1998, Schmitz et al., 1997, Vannuffel et al., 1995, Towner et al., 1998). Following this, meticillin resistance may then be detected either by phenotypic assays using chromogenic agar media and disc diffusion methods (Andrews, 2009, Perry et al., 2004), or with rapid molecular identification methods such as the Cepheid GenXpert (http://www.cepheid.com/systems-and-software/genexpert-system) and BD-GenoOhm systems (http://bd.com/geneohm/english/products/idi_mrsa.asp). These methods detect the presence of a SCCmec element integrated into the S. aureus genome at the 3’ end of the orfX gene (Brown et al., 2005, Malhotra-Kumar et al., 2008, Nathwani et al., 2008).

Adding a high resolution melt based rapid method (as described in the preceding chapter) as an adjunct to this process, may well improve speed of diagnosis of infection with PVL-positive S. aureus by allowing rapid detection of the PVL
encoding genes. Like other molecular techniques however, this method would be limited by a general lack of RT-PCR facilities in diagnostic settings and hence would be more suited for use in research or reference laboratories. In addition, this method has the same limitation as the PCR technique in that it simply involves detection of the encoding genes without providing any information regarding the actual expression of the gene. Therefore, a method which both confirms PVL production and gives some indication of the amount of toxin produced by different *S. aureus* isolates could potentially be useful not just clinically but also as a research tool.

By 2009, when the present work was initiated, a few studies had already attempted to quantify PVL production (Badiou et al., 2008, Hamilton et al., 2007, Tseng et al., 2009, Oishi et al., 2008). The majority of these studies (3/4), described enzyme-linked immunosorbent assays (ELISAs) for PVL detection and quantification using a polyclonal antibody specific to either the LukF (Hamilton et al., 2007, Tseng et al., 2009) or LukS (Badiou et al., 2008) PVL subunits. These ELISA protocols had however been designed as a research tool in quantifying PVL toxin production *in vitro* as part of a larger study, rather than for use specifically as a diagnostic tool in a clinical setting.

5.1.1 **Enzyme-Linked Immunosorbent Assay (ELISA)**

Immunnoassays which depend on the immunological reactions between an antigen and its specific antibody have been widely applied as a diagnostic tool in the clinical microbiology laboratory (Peruski and Peruski, 2003, Atchison et al., 2009, Tuke et al., 2008, Planche et al., 2008, Diggle and Clarke, 2006, Thorburn et al., 2004, Fry et al.,
2009, Andre et al., 2008, Trotter et al., 2010, Smith et al., 2009). These tests have been used in the diagnosis of infections caused by several bacteria and viruses, examples of which include: rotavirus, hepatitis C virus, respiratory syncytial virus, *Clostridium difficile*, *Neisseria meningitidis*, *Streptococcus pneumoniae* and *Bordetella pertussis*. The predominant immunoassay system in use is the enzyme-linked immunosorbent assay (ELISA).

ELISA, which was first reported in the early 1970s (Engvall and Perlmann, 1971, Engvall et al., 1971, Engvall and Perlmann, 1972), differs from other immunoassays in two key areas. These include the adsorption of one immunological binding component to a solid phase support and the presence of an enzyme-labelled component. Several variations of the ELISA process exist. The basic principle of this technique may however be illustrated using the simplest ELISA protocol, the direct ELISA. In direct ELISA (Figure 5.1), the test solution is first incubated with the solid phase support to allow antigen binding, followed by washing to remove unbound components. This is then followed by the addition of a conjugated enzyme-labelled antibody specific for the target antigen, and a second wash to again remove unbound components. Binding of the conjugated enzyme-labelled antibody (Ab′) takes place only if the test antigen (Ag) is bound to the solid phase, resulting in an Ag-Ab′ complex. In the absence of the test Ag, this enzyme-labelled antibody is removed with the second wash. The final step in the ELISA protocol is the addition of the enzyme-substrate. If the assay is positive (i.e. when the test Ag is present), this substrate is cleaved to produce a measurable colour change, proportionate to the amount of antigen bound to the solid phase support.
Figure 5.1: Illustration of the direct ELISA process.

In direct ELISA, the presence of the test antigen bound to the solid phase is detected via specific antigen-specific antibody conjugated to an enzyme (Ab'). The enzyme cleaves a specific substrate to produce a detectable coloured reaction.

Ag ( ); (Ab')-Enzyme conjugate ( ); Coloured product ( )

Ag: Antigen; Ab: Antibody
Several different solid phases such as beads (Stiffler-Rosenberg and Fey, 1978), tubes and plates (Voller et al., 1976) have been employed in ELISA. While the polystyrene tube was the predominant solid phase system used in initial studies (Engvall and Perlmann, 1971, Engvall et al., 1971, Engvall and Perlmann, 1972), at present, the polystyrene microtitre well plate is the more common solid phase format now in use. This format was first reported in 1974 (Voller et al., 1974) and has the advantage of being cheap and requiring only small volumes of reagents, hence reducing costs.

At present, there are four main variants of the ELISA technique. These four variants, the Direct, Indirect, Sandwich and Competition assays, all vary in complexity. Unlike the sandwich ELISA system (Figure 5.2 B) in which a specific primary antibody is the component bound to the solid surface, both direct and indirect ELISA first involve the binding of antigen in a test sample to the solid phase. The indirect ELISA (Figure 5.2 A) then further differs from the direct technique, by involving the sequential use of two antibodies (one unlabelled “primary” antibody and the second enzyme-labelled). The primary antibody recognises the test antigen. The second enzyme-labelled antibody specifically detects IgG from the animal species in which the primary antibody was produced (for example anti-rabbit IgG). Hence, this provides a greater degree of flexibility in the ELISA design. Indirect ELISA is therefore more commonly used in diagnosis (Crowther, 2000). Aspects of these three techniques are combined in the fourth variant of ELISA – competitive ELISA. This technique is so named because it involves competing the specific antigen present in a test sample against a labelled reference antigen of known concentration to produce a signal inversely proportional to the level of competing antigen present in the test sample (Anon, 1976, Voller et al., 1976).
Figure 5.2: Representations of Indirect (A) and Sandwich (B) ELISA processes

In indirect ELISA (A), the test antigen (Ag) bound to a solid phase is detected via a primary antibody (Ab) which is further detected via a secondary enzyme-labelled Ab (Ab'). The enzyme conjugated to the secondary Ab cleaves a specific substrate to produce a detectable coloured reaction. In Sandwich ELISA (B), the test Ag is detected via a primary capture Ab bound to a solid phase and a secondary enzyme-labelled detection Ab. This enzyme conjugated to the secondary Ab cleaves a specific substrate to produce a detectable coloured reaction.

Ag ( ); Ab ( ); Ab-Enzyme conjugate ( ); Coloured product ( )
Aims

The first aim of this chapter was to develop and validate an ELISA protocol which could possibly be applied in diagnosis. At the onset of the present study, no such standard system was commercially available for the detection and quantification of the PVL toxin in the clinical microbiology laboratory. In addition to this, a second aim in setting up this system was to apply it locally as a research tool to quantify toxin production in our clinical collection of PVL-positive isolates. A previous study (Sloan, 2010) had noted a wide variation in PVL toxin production among this group of isolates using Western blot analysis. In addition, it appeared in these initial studies, that strains producing high levels of PVL also produced high levels of HLA. Therefore the present study also aimed to further explore this observation by the development of a second ELISA to detect HLA so that production of both toxins by these strains could be quantified and compared.
5.2 Materials and Methods

5.2.1 Pre-protein expression and purification steps

5.2.1.1 Background

The first stage in the ELISA development involved production of recombinant LukF and S. aureus alpha toxin (HLA) proteins using the widely employed pET expression system (Sorensen and Mortensen, 2005) first described in the mid-1980s. This system makes use of a powerful T7 promoter carried on the pET vector and not found naturally in bacteria. High level expression of target genes from the T7 promoter takes place in the presence of the bacteriophage T7 RNA polymerase which is integrated in the genome of an E. coli host (Studier and Moffatt, 1986).

pET-21d is an engineered plasmid expression vector which contains a T7/lac hybrid promoter. This hybrid contains the molecular switch of the lac operon. Therefore, in the absence of an inducer, the T7/lac hybrid promoter is ‘switched off’ due to the presence of the constitutively expressed lactose repressor protein (LacI). This repressor protein shares the same binding site as the polymerase, hence binding of LacI prevents transcription (Wilson et al., 2007). Therefore for transcription to take place, an inducer must be present. Isopropyl-β-D-thiogalactopyranoside (IPTG) is a commonly used inducer. This binds the LacI protein, reducing its affinity for the operator site and enabling polymerase binding.

The T7 promoter is however not recognised by E. coli RNA polymerase. Therefore, a compatible expression host is also needed for transcription to take place. BL21/DE3 is
an engineered *E. coli* strain lysogenized by a DE3 phage fragment. The DE3 phage encodes the T7 RNA polymerase, making *E. coli* BL21/DE3 a suitable host for expression from pET-21d recombinant plasmids. Similar to the T7/lac promoter, these genes are under the control of an IPTG inducible *lacUV5* promoter.

Other salient features of the pET-21d expression vector include the presence of a gene encoding ampicillin resistance (allowing selection for plasmid uptake and maintenance in *E. coli*) and a sequence encoding a hexa-histidine tag (allowing rapid purification of tagged protein). Furthermore, *E. coli* BL21/DE3 lacks genes encoding ompT and Lon proteases in BL21/DE3 (to ensure the production of intact recombinant proteins) (Baneyx, 1999).

5.2.1.2 Construction of pET-21d(+)-*hla*H35L plasmid expression construct

pGEMT containing the *hla*-H35L gene (kindly provided by Dr Alan Cockayne) was first used as the template in a PCR designed to incorporate the *NcoI* and *XhoI* restriction sites into the gene fragment. Amplification was carried out with a forward 5′-ATAATCCATGGCAGATTCTGATATTAATTTAAAC-3′ and reverse 5′-CGATCTCGAGATTACTTCTTCTTTTTCTCCCA-3′ primer pair containing the *NcoI* and *XhoI* restriction sites (underlined) respectively. The reverse primer was designed to allow incorporation of a C′ terminal hexa-histidine tag on the expressed protein.
The *hla*-H35L gene used in the present study contains a mutation which causes a loss of haemolytic activity, resulting in reduced lethality to mice thereby making it useful in antibody generation (Menzies and Kernodle, 1994).

Following PCR amplification, the resulting fragment was A-tailed (2.11.1) and ligated (2.11.2) into pGEMT before transformation into *E. coli* Top10 and selection of ampicillin – resistant transformants. A plasmid (designated pKO1) was isolated from one transformant and the DNA sequence of the *hla*-H35L insert confirmed using M13F and M13R primers. (M13F: 5′-GTAAAACGACGGCCAGT-3′ and M13R: 5′-GGAAACAGCTATGACCATG-GGAAACAGCTATGACCATG-3′)

Following restriction enzyme digestion of both pKO1 and pET-21d(+) using NcoI and *XhoI* (2.11.3), and gel purification of the vector and insert fragments, the *hla*-H35L gene was ligated into pET-21d(+) (Novagen) (Figure 5.3). Ligation mixtures were electroporated into *E. coli* Top10 and ampicillin – resistant transformants selected. A plasmid (designated pKO2) was isolated from one transformant, sequenced and the insertion site of the *hla*-H35L insert confirmed using the T7 promoter primer (5′-TAATACGACTCACTATAGGG-3′).

5.2.1.3 pET-21d(+)-*lukF* expression construct

A pET-21d(+) construct (pKO3) containing the complete *lukF* gene (Okolie, 2009) was kindly provided by Dr. Charles Okolie. LukF expressed from this construct has a C′ terminal hexa-histidine tag.
Relevant features of the pET-21d(+) plasmid (Novagen) include: 1. An inducible T7/lac hybrid promoter (circled red). Transcription from this promoter in a compatible host takes place in the presence of an inducer (IPTG). The inducer binds the LacI repressor protein enabling polymerase binding while the compatible host (e.g. *E. coli* BL21/DE3) carries the genes encoding the polymerase specific for the T7 promoter (T7 RNA polymerase). The promoter region also includes the DNA sequence encoding for a C' terminal hexa-histidine tag. 2. Multi-cloning sites (circled blue) that allow the in-frame cloning of a gene or gene fragment downstream of the T7/lac promoter. 3. Two labelled open reading frames (ORF1 and ORF2). These contain the gene which confers ampicillin resistance (circled black – ORF 1) and the *lacI* gene (circled yellow – ORF 2) respectively.
5.2.2 **Protein expression**

The next step in this study was the expression of recombinant LukF and HLA-H35L proteins. For protein expression, both pET-21d constructs (pKO2 and pKO3) were first transformed separately into *E. coli* BL21/DE3 cells as described in 2.11.5. Transformants (BL21/DE3pKO2 and BL21/DE3pKO3) were subsequently selected on LB agar containing 100 µg/ml of ampicillin.

Single transformant colonies of BL21/DE3pKO2 and BL21/DE3pKO3 were inoculated into LB broth supplemented with 100 µg/ml ampicillin and grown with shaking overnight at 37°C, in a G25 shaker incubator (New Brunswick Scientific Co Inc. Edison, New Jersey USA). Overnight cultures were then diluted 1:100 into fresh LB broth supplemented with 100 µg/ml ampicillin and grown with shaking at 200 rpm at 37°C until an OD$_{600nm}$ of 0.4 was achieved.

Expression of both the HLA-H35L and LukF recombinant proteins was then induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and growth continued with shaking at 37°C for 2 h. Following this, cells were harvested by centrifugation at 9,000 rpm for 10 min at 4°C (Allegra X-22R, Beckman Coulter).

An aliquot of the harvested cells was subsequently analysed in order to confirm the presence, relative amount and solubility of the expressed proteins. To do this, cells resuspended in PBS were disrupted by sonication on ice (2.12.1), fractionated (2.12.2) by centrifuging at 13,000 rpm for 10 min, and the soluble and insoluble fractions run
on an 11.5% SDS-PAGE gel and proteins visualised using Coomassie blue stain (2.12.4).

5.2.3 **Protein Purification**

Following a large scale expression of proteins using 500 ml cultures under optimised conditions, the expressed proteins were initially purified using a Novagen His Bind Resin® protein purification kit according to the manufacturer’s instructions. In this process, purification depends on the ability of His-tagged proteins to bind nickel ions in the resin. Unbound proteins are then removed in a subsequent wash and the recombinant His-tagged proteins eluted following displacement in buffer containing a high concentration of imidazole.

Purified proteins were subsequently dialysed using a Spectra/Por 1 dialysis tubing (6 – 8000 Da MWCO) into PBS to reduce the imidazole concentration, fractionated and the final purified protein stored in aliquots at -80°C.

5.2.4 **Protein Quantification**

As known concentrations of protein are required for ELISA development, the proteins were quantified using the Pierce® BCA protein assay kit according to manufacturer’s instructions. In brief, 0.1 ml of test sample was added to 2 ml of BCA working reagent. Protein standards containing known concentrations (0 µg/ml – 2000 µg/ml) of bovine serum albumin (BSA) were tested simultaneously for the preparation of a standard curve. Following 30 min incubation at 37°C, samples were cooled to room
temperature and absorbance readings determined at 562 nm. Protein concentrations of unknown samples were then determined from the standard curve.

5.2.5 Protein Biotinylation

As competitive ELISA depends on the competition between a labelled and unlabelled version of the same protein, samples of both the HLA-H35L and LukF antigens used in the present study were biotinylated using EZ-Link® NHS-LC-Biotin (Thermo Scientific, Surrey UK) as a first step in ELISA development. This process was carried out in a glass bijoux according to manufacturer’s instructions and involved incubating 25 µl of a freshly prepared 2 mg/ml biotin solution (dissolved in DMF) with 1 mg of protein (in PBS) at 4°C for 1 h. Following this, unbound excess biotin was removed by overnight dialysis into PBS using a Spectra/Por 1 dialysis tubing (6 – 8000 Da MWCO). The biotinylated proteins were stored in aliquots at -20°C.

5.2.5.1 ELISA to evaluate protein biotinylation

In order to evaluate the extent of protein biotinylation and to establish appropriate working concentrations, an ELISA was set up as illustrated in Figure 5.4. Mouse monoclonal antibodies specific for either HLA-H35L or LukF used in the assays were supplied by Prof. Lyndy Durrant, Oncology Unit, Nottingham City Hospital Campus, University of Nottingham.

A 96-well Maxisorp microtiter plate (Nunc, Roskilde, Denmark) was coated overnight at 4°C with 50 µl per well of appropriate antibody against each protein (5 µg/ml in PBS). Exposed well surfaces were subsequently blocked by the addition of 100 µl of
3% w/v BSA in PBS for 1 h on ice. The plates were then washed three times using PBS containing 0.1% v/v Tween-20 (PBST). The next step involved the addition to triplicate wells of 50 µl of varying concentrations of biotinylated protein (10, 3, 1, 0.3, 0.1, 0.03, 0 µg/ml) and incubation for 1 h on ice. The plates were then washed three times with PBST. 50 µl of 1 µg/ml Pierce® high sensitivity streptavidin-horseradish peroxidase (HRP) conjugate in PBS was added to each well and the plates were then incubated on ice for 1 h. The plates were then washed six times with PBST before addition of 150 µl/well of ready to use 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) [ABTS] (Thermo Scientific, Surrey, UK) and incubation for 45 min at room temperature. The absorbance of each well was then read at 405 nm on a Gen5 microplate reader (Biotek, Bedfordshire, UK).

5.2.6 Competitive ELISA

Subsequent to the successful labelling of the HLA-H35L and LukF proteins, a competitive ELISA was set up as previously described (Kurta, 1985). The protocol for this is as described in Section 5.2.5.1, with one additional step (the competition step proper). This involved addition of 25 µl of various concentrations of competing non-biotinylated test protein, test sample or buffer control to appropriate microtitre plate wells and incubation for 1 h on ice prior to the addition of biotinylated protein (25 µl) with no wash in between these additions (Figure 5.5).
Figure 5.4: Flowchart depicting the two major ELISA protocols used in the present study.

The flowchart shows the sequential addition of the ELISA components separated by an incubation and wash step. The competitive ELISA protocol differs from the biotinylation testing ELISA protocol by incorporating an extra step involving the addition of unlabelled protein, test sample or buffer control without any additional wash step (as indicated by X), prior to addition of biotinylated protein.

ABTS: 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
5.3 Results

5.3.1 Confirmation of hla-H35L sequence prior to protein expression

Before cloning of the mutated hla-H35L gene into the pET-21d(+) plasmid expression vector, sequencing of the hla-H35L gene in pKO1 and subsequent comparison against the S. aureus Newman hla gene (NWMN_1073, Pubmed accession number AP009351.1) confirmed the presence of the H35L protein mutation and the successful insertion of the restriction sites as indicated below (Figure 5.5). A second C to T SNP at nucleotide 453 was however found in the mutant hla gene but this mutation was silent, not resulting in a second amino acid change (Figure 5.6).
Figure 5.5: DNA sequence alignment of *hla*-H35L in pKO1 and the *S. aureus* Newman *hla*

The DNA sequence alignment shows the position of the H35L mutation (highlighted in red) and a second silent SNP (highlighted in blue). The *NcoI* and *XhoI* restriction sites are highlighted in yellow.
Figure 5.6: Protein sequence alignment of the HLA_NWMN and HLA-H35L protein sequences.

The protein sequence alignment shows the predicted single amino acid difference (highlighted in red) between the HLA-H35L protein sequence obtained from the gene sequence described in Figure 5.5 and the S. aureus Newman hla gene (NWMN_1073, Pubmed accession number AP009351.1). Both proteins are ≈33.2 kDa
5.3.2 **Pilot studies to confirm HLA-H35L and LukF protein expression**

SDS-PAGE analysis was used to confirm the expression of the ≈33.2 kDa recombinant HLA-H35L protein containing a C-terminal hexa-histidine tag in BL21/DE3pKO2 following a 2 h induction with 1 mM IPTG (Figure 5.7, Lane 2). Analysis of cell fractions prepared as described in Section 2.12.4 showed that despite the presence of some HLA-H35L in the insoluble fraction (Lane 4), the amount of protein in the soluble, cytoplasmic fraction (Lane 3) was considered sufficient to attempt purification using Ni affinity chromatography.

For the ≈34 kDa recombinant LukF protein on the other hand, although protein expression at 37°C was quite efficient (Figure 5.8, Lane 1), most of the protein was insoluble (Lane 3). In this case, it was therefore necessary to optimise the protein expression process for soluble protein production. Growth temperature has previously been reported to affect protein solubility (Klein and Dhurjati, 1995). Expression of the recombinant LukF protein was therefore carried out at both 30°C and 37°C to ascertain if this improved the production of soluble forms of the protein. This process showed that while the yield of protein following induction at 30°C (Lane 5) was slightly less than the total amount of protein produced at 37°C, induction at 30°C resulted in more LukF protein present in the soluble fraction (Lane 6) as opposed to the insoluble fraction. Consequently induction at 30°C was used for all further expression experiments with LukF.
Figure 5.7: SDS-PAGE analysis of recombinant HLA-H35L protein expression at 37°C in *E. coli* BL21/DE3pKO2.

The expression of recombinant HLA-H35L from *E. coli* BL21/DE3pKO2 cells following induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 37°C was carried out as described in Section 5.2.2. The separation and visualisation of the protein fractions were carried out as described in Section 2.12.

Lane A: Protein molecular weight standards; Lane 1: Total proteins of uninduced *E. coli* BL21/DE3pKO2 cells; Lane 2: Total proteins of IPTG induced BL21/DE3pKO2 cells; Lane 3: Soluble fraction of IPTG induced BL21/DE3pKO2 cells; Lane 4: Insoluble fraction of IPTG induced BL21/DE3pKO2 cells.
Figure 5.8: SDS-PAGE analysis of recombinant LukF protein expression at 30°C or 37°C in *E. coli* BL21/DE3pKO3.

The expression of recombinant LukF protein from *E. coli* BL21/DE3pKO3 cells following induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was carried out at both 30°C and 37°C as described in Section 5.2.2. The separation and visualisation of the protein fractions were carried out as described in Section 2.12.

Lane A: Protein molecular weight standards; Lane 1: Total proteins of IPTG induced BL21/DE3pKO3 cells grown at 37°C; Lane 2: Soluble fraction of IPTG induced BL21/DE3pKO3 cells grown at 37°C; Lane 3: Insoluble fraction of IPTG induced BL21/DE3pKO3 cells grown at 37°C; Lane 4: Total proteins of uninduced BL21/DE3pKO3 cells; Lane 5: Total proteins of IPTG induced BL21/DE3pKO3 cells grown at 30°C; Lane 6: Soluble fraction of IPTG induced BL21/DE3pKO3 cells grown at 30°C; Lane 7: Insoluble fraction of IPTG induced BL21/DE3pKO3 cells grown at 30°C.
5.3.3 Affinity Protein Purification

Purification using the Novagen His Bind Resin® kit, from soluble fractions prepared from 500 ml IPTG-induced cultures was more successful for the LukF protein than the HLA-H35L protein (Figure 5.9). For both of these proteins monitoring of the purification process using SDS-PAGE revealed efficient selective binding of the His tagged protein. This was indicated by the absence of an appropriate sized protein in the column run through compared to the starting sample (Figure 5.9; Lanes 1 and 2). Neither protein was eluted from the column in the wash buffer (Lane 3).

However, following protein elution by displacement with 25 mM imidazole, while the process for the LukF protein produced a high yield (1800 µg/ml) with a low level of contaminating protein, the level of contaminating protein after the HLA-H35L purification process was deemed unacceptable (Lanes 4 – 9). This suggested that further modification of the purification process was required for the HLA-H35L protein.
Figure 5.9: SDS-PAGE analysis of recombinant LukF (top panel) and HLA-H35L (bottom panel) expressed in *E. coli* BL21/DE3 and purified using the Novagen His Bind Resin.

The expression of both recombinant HLA-H35L and LukF proteins from *E. coli* BL21/DE3 cells following induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was carried out as described in Section 5.2.2 and expressed recombinant proteins purified using the Novagen His Bind Resin as described in Section 5.2.3. The separation and visualisation of the protein fractions were carried out as described in Section 2.12.

Lane A: Protein molecular weight standards; Lane 1: Resuspended IPTG induced cell pellets (Starting Material); Lane 2: Initial binding buffer run through (purification flow through following initial binding); Lane 3: Wash buffer run through; Lanes 4 – 9: Sequential fractions collected during protein elution.
5.3.4 Modification of HLA-H35L purification procedure

As protein elution using the Novagen His Bind Resin® protein purification kit proved rather unsatisfactory for the HLA-H35L protein due to the presence of many contaminating proteins (Figure 5.9), a different purification system was explored. The HLA-H35L protein was therefore purified using His Trap affinity columns on the BioRad® Biologic LP affinity chromatographic system. To achieve this, after induction with IPTG, BL21/DE3pKO2 cell pellets were resuspended in NiC buffer (20 mM imidazole, 50 mM NaCl, 10% v/v glycerol), lysed by sonication (2.12.1), and cell debris removed by both centrifugation at 13,000 rpm for 30 min at 4°C and filtration using a 0.2 µm filter (Sartorius Biotech, Germany). The resulting cell-free supernatant was then applied to the column charged with 50 mM NiSO₄ at a flow rate of 1 ml/min. After washing the column with buffer A (50 mM NaCl in PBS) to remove unbound proteins, the His-Tagged protein was eluted from the column using a linear gradient of imidazole (50 → 400 mM) in buffer A and the eluted fractions were collected and examined using SDS-PAGE.

This approach resulted in a more efficient purification of the HLA-H35L protein (Figure 5.10, Lanes 9 – 17), with protein elution occurring over a wide range of imidazole concentrations and resulting in a relatively pure product.

Dialysis of the pooled protein fractions (10 – 17) against PBS overnight to remove imidazole resulted in precipitation of approximately 55% of the HLA-H35L protein leaving about 45% of protein in solution (Figure 5.11, Lane 3). This protein had an apparent molecular mass in agreement with the expected value of approximately
33.2 kDa. The lack of haemolytic activity of the soluble purified mutated HLA-H35L protein was confirmed by assessing haemolysis on horse blood agar with an equal concentration of purified S. aureus alpha toxin (Sigma-Aldrich, Dorset UK) as a positive control (Appendix 4). The purified, soluble protein (190 µg/ml) was stored in aliquots at -80°C.
Figure 5.10: SDS-PAGE analysis of recombinant HLA-H35L expressed in *E. coli* BL21/DE3 and purified using the BioRad® Biologic LP affinity chromatographic system

The expression of recombinant HLA-H35L protein from *E. coli* BL21/DE3pKO2 cells following induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was carried out as described in Section 5.2.2 and expressed recombinant proteins purified using the BioRad® Biologic LP affinity chromatographic system as described in Section 5.3.4. The separation and visualisation of the protein fractions were carried out as described in Section 2.12.

Lane A: Protein molecular weight standards; Lane 1: Soluble fraction of IPTG induced BL21/DE3pKO2 cell pellet resuspended and lysed in NiC buffer (Starting Material); Lane 2: Initial binding buffer run through; Lanes 3 – 17: Sequential fractions collected during protein elution.
Figure 5.11: SDS-PAGE analysis of purified and dialysed recombinant HLA-H35L.

The expression of recombinant HLA-H35L protein from E. coli BL21/DE3 cells following induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was carried out as described in Section 5.2.2 and expressed recombinant proteins purified using the BioRad® Biologic LP affinity chromatographic system as described in Section 5.3.4. The separation and visualisation of the protein fractions were carried out as described in Section 2.12. The dialysis of the purified protein was carried out as described in Section 5.2.3.

Lane A: Protein molecular weight standards; Lane 1: Pooled, eluted purified HLA-H35L recombinant protein prior to dialysis; Lane 2: Soluble fraction of HLA-H35L recombinant protein after dialysis against PBS; Lane 3: Insoluble fraction of HLA-H35L recombinant protein after dialysis against PBS.
5.3.5 Use of purified recombinant LukF and HLA-H35L proteins in competitive ELISA development

5.3.5.1 Protein Biotinylation

Protein biotinylation is a necessary step in the development of a competitive ELISA. An assessment of the biotinylation process (Section 5.2.5) which was carried out as described in Section 5.2.5.1 indicated a successful biotinylation of both purified recombinant LukF and HLA-H35L proteins (Figure 5.12). In this process, addition of 50 µl of increasing amounts of biotinylated proteins (10, 3, 1, 0.3, 0.1, 0.03, 0 µg/ml) to the microtitre wells coated with the relevant monoclonal antibodies, resulted in concurrent increase in OD readings (Figure 5.12) following addition of streptavidin-HRP conjugate and the HRP substrate, ABTS.

Results of the biotinylated HLA-H35L protein assay however, appeared to indicate a less efficient biotinylation process (Figure 5.12 A). Results of the biotinylated LukF assay generated OD$_{405}$ values ranging from 0.62 to 2.98 and resulting in a maximum 4.8 fold difference between lowest and highest concentrations tested. In contrast, the HLA-H35L assay resulted in values which ranged only from 0.46 to 0.83 (maximum fold difference of 1.8). In addition, use of the higher concentrations of biotinylated HLA-H35L (10 µg/ml and 3 µg/ml) appeared to have a saturation/inhibition effect. These resulted in slightly lower OD$_{405}$ values (0.803 and 0.763 respectively) than those obtained with addition of 1 µg/ml protein (0.828) (Figure 5.12 A).

Due to the less efficient biotinylation achieved using the HLA-H35L protein, a second biotinylation was carried out on both LukF and HLA-H35L in an attempt to improve
efficiency and to investigate reproducibility of the reaction. This was done using the same process as described in Section 5.2.5 but using 75 µl of 4 mg/ml EZ-Link® NHS-LC-Biotin per 1 mg of protein. When used in ELISAs, proteins from the second biotinylation provided a higher level of discrimination resulting in OD values ranging from 0.62 to 2.98 and 0.07 to 1.35 with concurrent maximum fold differences of 4.81 and 19.29, for the HLA-H35L (Figure 5.12 C) and LukF proteins (Figure 5.12 D) respectively.

Based on these results for both assays, a working concentration of 3 µg/ml biotinylated protein was subsequently selected for use in further competition assays. This concentration generated fold differences of 4.1 and 7.4 for OD values in the presence and absence of HLA-H35L and LukF proteins respectively. These differences were considered to be sufficient to be able to assess competitive inhibition using test samples.
Figure 5.12: Determination of the biotinylation efficiency of the recombinant HLA-H35L and LukF proteins obtained in two separate biotinylation reactions.

The ELISA protocol used to assess the efficiency of the biotinylation process was carried out as described Section 5.2.5.1. The 1\textsuperscript{st} biotinylation reactions (A and B) were carried out using 25 µl 2 mg/ml of EZ-Link\textregistered NHS-LC-Biotin. The 2\textsuperscript{nd} biotinylation reactions (C and D) were carried out using 75 µl 4 mg/ml of EZ-Link\textregistered NHS-LC-Biotin. The optical density readings reflect the colour change produced from the breakdown of ABTS bound to streptavidin-HRP conjugate which allows the detection and quantification of biotinylated protein present in the assay.
5.3.6  Competitive ELISA exploring competitive ability of biotinylated proteins

Following the assessment of protein biotinylation, an initial competitive ELISA was carried out using known concentrations of unlabelled HLA-H35L and LukF in PBS as described in Section 5.2.6.1. These initial ELISA tests were designed to explore the ability of unlabelled protein to inhibit binding of biotinylated protein using standard protein concentrations of 100, 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01, 0.003 and 0 µg/ml. Both batches of biotinylated HLA-H35L and LukF (Section 5.3.5.1), were used in these assays which clearly demonstrated the competitive abilities of the unlabelled proteins. However, comparing results using both sets of biotinylated proteins highlighted the effect of poor biotinylation on experimental outcomes. Batch 1 biotinylated proteins (Figure 5.13 A and B) provided a low level of discrimination between the lowest and highest concentrations of non-biotinylated protein tested. This was as indicated by maximum OD differences of 0.33 and 0.55 for HLA-H35L and LukF respectively as opposed to the 11.1 and 4.5 fold OD differences noted with the batch 2 biotinylated proteins (Figure 5.13 C and D).

The degree of competition shown in Figure 5.13 C and D, was also expressed quantitatively as percentage inhibition (Figure 5.13 E and F) calculated using the formula \[ 1 - (B/B_0) \] × 100.
Quantitative representation of level of competition

Figure 5.13: Pilot competitive ELISA results comparing the ability of two batches of biotinylated HLA-H35L and LukF to inhibit binding of unlabelled proteins.

The competitive ELISA protocol used to explore the ability of unlabelled protein to inhibit binding of biotinylated protein was carried out as described in Section 5.2.6. The charts show variations in OD readings obtained after addition of varying concentrations of unlabelled proteins to assay wells, in the presence of a fixed amount of biotinylated HLA-H35L or LukF.

Panels A and B show the results from competitive assays using batch 1 HLA-H35L and LukF proteins respectively.
Panels C and D show the results from competitive assays using batch 2 HLA-H35L and LukF proteins respectively.
Panels E and F show results from C and D respectively, expressed as percent (%) inhibition by the different concentrations of competing protein relative to uncompeted control.
5.3.7 Pilot studies using spiked culture media in competitive ELISA

Three culture media (BHI, LB and CYGP) were spiked with known concentrations of unlabelled LukF or HLA-H35L and used in the competition ELISA. Lower OD values were observed in the presence of BHI and LB culture media spiked with protein as opposed to assays using culture media alone (Figure 5.14 A and B) clearly indicating inhibition of biotinylated protein binding.

The data generated using CYGP medium however, identified a potential problem with the interpretation of the ELISA results. CYGP is the medium of choice for S. aureus protein studies (Novick, 1991). For the HLA-H35L assay, addition of CYGP medium alone competed with binding of the biotinylated protein (Figure 5.14 A). In the case of the LukF assay, the CYGP medium alone did not appear to cause interference. Addition of unlabelled LukF however failed to result in any inhibition (Figure 5.14 B). Rather, higher OD values of 1.73 and 1.59 were obtained with 1 µg/ml and 10 µg/ml of LukF protein added respectively as opposed to an OD value of 1.36 obtained with addition of CYGP alone.
Figure 5.14: Results of competitive ELISA using various culture media (BHI, LB and CYGP) spiked with HLA-H35L or LukF
Charts show OD values generated following a competitive ELISA carried out as described in Section 5.2.7 using three different culture media (BHI, LB and CYGP) spiked with the indicated protein concentrations.

HLA-H35L or LukF diluted in PBS or PBS alone served as the controls

LB: Luria Bertani broth; BHI: Brain heart infusion broth; CYGP: Casamino acids-yeast extract-glycerophosphate broth; PBS: Phosphate buffered saline.
Based on analysis of the components of the various media, this phenomenon was suspected to be due to the β-glycerophosphate component of the CYGP medium. Therefore, a competitive ELISA was carried out to explore this using dilutions of CYGP (1:10 and 1:100 in PBS), with and without the β-glycerophosphate component. All dilutions were spiked with identical amounts of unlabelled protein (10 µg/ml). Data shown in Figure 5.15 A and B show that removal of the β-glycerophosphate component or dilution of the CYGP broth in PBS prior addition of unlabelled protein and use in the ELISA, resulted in clear inhibition of binding of the biotinylated proteins.

For the HLA-H35L competitive ELISA, in contrast to an OD value of 0.36 obtained following addition of CYGP medium alone (i.e. without protein), addition of CYGP minus β-glycerophosphate (CYGP-) alone (i.e. without protein), resulted in an OD value of 1.88. This was similar to the OD value of 1.82 obtained with addition of the PBS control. In addition, diluting the complete CYGP medium in PBS prior to spiking with 10 µg/ml of protein produced identical optical density readings of 0.23 and 0.23 for the 1:10 and 1:100 dilutions respectively as opposed to an optical density of 0.57 obtained when undiluted CYGP was spiked with the same amount of protein.

For the LukF competitive ELISA, addition of CYGP medium alone (i.e. without protein) also showed some inhibition of biotinylated protein binding resulting in an OD value of 1.02 which was lower than obtained with the PBS control (1.36). Unexpectedly however, no additional competition was observed on addition of unlabelled protein. Addition of CYGP containing 10 µg/ml of LukF actually
resulted in an increased OD value of 1.68. These anomalies were resolved by either leaving the β-glycerophosphate component out of the medium or by dilution of the CYGP broth (Figure 5.15 B). These assays gave OD values similar to those obtained using the PBS control spiked with 10 µg/ml of the protein.
Figure 5.15: Effects of β-glycerophosphate and CYGP dilution on competitive ELISA results.

Charts comparing OD values generated following competitive ELISA carried out as described in Section 5.2.6 using various dilutions of CYGP with and without β-glycerophosphate, spiked with 10 µg/ml HLA-H35L or LukF.

HLA-H35L or LukF diluted in PBS or PBS alone served as the controls

CYGP: Casamino acids-yeast extract-glycerophosphate broth; PBS: Phosphate buffered saline.
CYGP-: CYGP broth without β-glycerophosphate.
To determine if other *S. aureus* proteins present in culture supernatants could interfere with the performance of the LukF competitive ELISA, supernatants from three known PVL-negative isolates (TS3, TS4, TS10) were tested in the assay. Three dilutions of filter sterilised supernatants of overnight CYGP broth cultures (1:2, 1:5 and 1:10 in PBS) were tested. In comparison with the positive and negative controls, the data presented in Figure 5.16 confirm that other proteins present in these culture supernatants did not interfere with the performance of the LukF assay. All three dilutions of the PVL-negative test isolates resulted in OD values ranging from 1.15 – 1.24. These values were similar to the OD value of 1.35 obtained in the no competitor control.
Figure 5.16: Specificity testing of LukF ELISA using CYGP culture supernatants of PVL-negative *S. aureus* isolates

Charts show OD values generated following competitive ELISA carried out as described in **Section 5.2.6** using CYGP culture supernatants of PVL-negative *S. aureus* isolates.

PBS alone or PBS containing the indicated LukF concentrations served as the controls.

Test isolates represent a random selection of PCR confirmed PVL-negative isolates in our collection of strains obtained from the NUH trust between 2008 and 2009.
5.3.8 **Validation of ELISA**

5.3.8.1 **Generation of Standard Inhibition Curves for protein quantification using the competitive ELISAs**

In order to obtain quantitative data of concentrations of HLA or LukF in culture supernatants using the competitive ELISAs, it was first necessary to generate standard inhibition/competition curves using recombinant HLA-H35L or LukF. The non-linear competition curves ([Figure 5.17](#)) were obtained by creating a plot of normalised signal \((B/B_0)\) against standard protein concentration using the GraphPad Prism 5.04 software. B values represent the mean absorbance in the presence of test protein while \(B_0\) is the mean absorbance of control biotinylated protein only wells (i.e. wells with no competing protein). Competition curve data was obtained from four independent experiments carried out in triplicate. OD values were obtained from competitive ELISAs as described in [Section 5.2.6](#), using known concentrations \((100, 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01, 0.003, 0 \mu g/ml)\) of either recombinant HLA-H35L or LukF proteins.

The HLA-H35L and LukF standard inhibition curves produced different dynamic ranges. LukF had a range of 0.04 – 0.54 \(\mu g/ml\) while HLA-H35L had a range of 0.01 – 0.14 \(\mu g/ml\). The limit of detection (LOD) was 0.007 \(\mu g/ml\) and 0.004 \(\mu g/ml\) for the LukF and HLA-H35L assays, respectively. Dynamic range which was calculated using the ‘log(Agonist) vs. response -- Find ECanything’ function of the Graph Prism 5.04 software, is defined as the protein concentration providing 20 – 80\% inhibition (IC 20 – IC 80 values) of the maximum signal. These values refer to the range of antigen concentrations in a specific assay at which accurate quantification is possible.
This is highlighted in Table 5.1 which shows the relationship between interpolated protein concentration based on normalised signal values, actual starting protein concentrations and recovery rate which was determined as previously described (Djoba Siawaya et al., 2008). For the LukF assay, higher concordant values were obtained for protein concentrations close to the dynamic range (0.1 – 1.0 µg/ml). Recoveries of the spiked LukF sample concentrations ranged from 70% to 120% within this range, in contrast concentrations outside this range resulted in recoveries ranging from 0% to 173% of the starting concentration. In the case of the HLA-H35L assay, this concordance was noted with concentrations ranging from 0.01 – 0.1 µg/ml. These concentrations yielded recoveries ranging from 67% to 119% while concentrations outside this range yielded 18% to 119% recoveries. Acceptable recovery ranges have been reported as 80% to 120% (Evan L. Chiswick, 2012, Giraudi et al., 1999) and 70% to 130% (Djoba Siawaya et al., 2008, Zhu-Shimoni et al., 2009).

The LOD is the lowest concentration of protein that can be reliably detected and was defined as the protein concentration providing a 10% inhibition (IC 10) of the maximum signal (Campas et al., 2007).
Figure 5.17: Standard inhibition calibration curves for HLA-H35L and LukF proteins generated using GraphPad Prism.

The data used to create the calibration curves were generated from four individual assays carried out in triplicates as described in Section 5.3.8.1. The standard calibration curves were obtained by plotting normalised signal (B/B₀) against protein concentration.

Vertical bars indicate the ±SD
Table 5.1: Correlation between actual and interpolated protein concentrations

<table>
<thead>
<tr>
<th>Actual Starting Protein Concentration (µg/ml)*</th>
<th>Interpolated Protein Concentration (µg/ml)**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LukF</td>
</tr>
<tr>
<td>100</td>
<td>115 [115%]</td>
</tr>
<tr>
<td>30</td>
<td>17.6 [59%]</td>
</tr>
<tr>
<td>10</td>
<td>14.3 [143%]</td>
</tr>
<tr>
<td>3</td>
<td>5.2 [173%]</td>
</tr>
<tr>
<td>1</td>
<td>1.2 [120%]</td>
</tr>
<tr>
<td>0.3</td>
<td>0.30 [100%]</td>
</tr>
<tr>
<td>0.1</td>
<td>0.07 [70%]</td>
</tr>
<tr>
<td>0.03</td>
<td>0.04 [133%]</td>
</tr>
<tr>
<td>0.01</td>
<td>0.014 [140%]</td>
</tr>
<tr>
<td>0.003</td>
<td>0.0025 [82%]</td>
</tr>
<tr>
<td>0.001</td>
<td>0.00005 [0%]</td>
</tr>
<tr>
<td>0</td>
<td>0.00005</td>
</tr>
</tbody>
</table>

The data presented in this table was generated as described in Section 5.3.8.1. The figures in bracket show the agreement between interpolated and actual protein concentrations. These values were calculated using the following formula:

\[
\text{Interpolated Concentration} - \text{Interpolated Concentration for } 0 \, \mu\text{g/ml competing protein} \times 100\% \\
\text{Actual Starting Concentration}
\]

*Actual starting protein concentration indicates known protein concentrations employed in the competitive ELISAs used to create the standard inhibition curve (Section 5.3.8.1)

**Interpolated protein concentration indicates the corresponding protein concentration for each standard determined from the standard inhibition curves.

***ND (Not Determined) indicates an inability to ascertain protein concentrations from the standard inhibition curves.

Concentrations around the dynamic range (shaded yellow) for both proteins resulted in a better agreement between the actual and interpolated protein concentration.
5.3.8.2 Application of Competitive ELISA to a panel of S. aureus isolates.

The developed competitive ELISAs were subsequently used to detect and quantify both PVL and HLA levels in overnight CYGP broth cultures of all the PVL-positive S. aureus isolates in our collection (Table 2.1 and Table 2.2). Two known PVL-negative isolates (RSS043 and RSS063) were also included in the analysis. Due to the number of test isolates (40), the assays were carried out in two batches (Set 1 and Set 2).

These assays revealed a wide variation in concentrations of LukF and HLA detected (Table 5.2). The concentrations ranged from 0.01 – 7.91 µg/ml for the LukF protein and 0.03 – 8.52 µg/ml for the HLA protein (Raw data is presented in Appendix 5). No LukF protein was however detected in 7.9% (3/38) of the PVL-positive test isolates or in the two PVL-negative isolates. This finding of a lack of detection of the LukF protein in PCR confirmed PVL-positive test isolates was in contrast to previous studies which reported a 100% detection of PVL protein expression in all strains harbouring the encoding genes (Loughman et al., 2009, Hamilton et al., 2007, Said-Salim et al., 2005, Badiou et al., 2010).

Based on a previous report of a possible positive correlation between the expression levels of LukF and HLA (Sloan, 2010) which could potentially have some significance with relation to variation in strain pathogenicity, the concentrations of the two proteins produced were compared for each isolates. Similar concentrations of both proteins (defined as concentrations differing by approximately 2 fold or less) were
produced by 28.9% (11/38) of isolates (Table 5.3) with concentrations of both LukF and HLA protein detected in TS24 as high as 7.91 µg/ml and 5.06 µg/ml respectively as compared to the lowest detectable protein concentration of 0.01 µg/ml (LukF, TS7). This phenomenon was however not universal among the isolates. For example the concentration of LukF protein detected for TS14 (1.11 µg/ml) was 22 fold more than that of HLA (0.05 µg/ml).
Table 5.2: Detection and Quantification of LukF and HLA proteins using the newly developed ELISA platforms

<table>
<thead>
<tr>
<th>Conc (µg/ml)</th>
<th>Not Detected</th>
<th>0 – 0.09</th>
<th>0.1 – 0.99</th>
<th>1.0 – 4.9</th>
<th>5.0 – 10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SET 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LukF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01 – 0.08)</td>
<td></td>
<td>(0.15 – 0.97)</td>
<td></td>
<td>(1.27 – 4.34)</td>
</tr>
<tr>
<td></td>
<td>TS7, TS12, TS15, TS5, TS13</td>
<td></td>
<td>TS9, TS1, TS2, TS6, TS19</td>
<td></td>
<td>TS23, TS20, TS18, TS14,</td>
</tr>
<tr>
<td></td>
<td>TS21, TS8, TS25, TS16</td>
<td></td>
<td>(ST22, 1518, 772, 22, 22)</td>
<td></td>
<td>TS17</td>
</tr>
<tr>
<td></td>
<td>(ST 30, 30, 30, 772, 22, 88, 772, 1, 30)</td>
<td></td>
<td></td>
<td></td>
<td>(ST22)</td>
</tr>
<tr>
<td><strong>HLA</strong></td>
<td></td>
<td></td>
<td>(0.05 – 0.07)</td>
<td></td>
<td>(1.33 – 4.62)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TS14, TS12, TS7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(ST22, 30, 30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SET 2</strong></td>
<td></td>
<td></td>
<td>(0.04 – 0.10)</td>
<td></td>
<td>(1.58)</td>
</tr>
<tr>
<td><strong>LukF</strong></td>
<td></td>
<td></td>
<td>(0.11 – 0.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TS29, TS30,</td>
<td></td>
<td>TS27, TS26, NRS158, NRS227,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NRS229, RSS043,</td>
<td></td>
<td>NRS157</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RSS063</td>
<td></td>
<td>(ST8, 1, 8, 25, 22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ST8, 8, 1, 22, 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HLA</strong></td>
<td></td>
<td>(0.03)</td>
<td>(0.11 – 0.87)</td>
<td>(6.70 – 8.52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TS29, TS30,</td>
<td></td>
<td>NRS194, NRS248, TS28,</td>
<td></td>
<td>RSS043,</td>
</tr>
<tr>
<td></td>
<td>NRS123, NRS162,</td>
<td></td>
<td>NRS157, NRS229, NRS192,</td>
<td></td>
<td>NRS227</td>
</tr>
<tr>
<td></td>
<td>NRS185, RSS289,</td>
<td></td>
<td>TS27, RSS063, NRS255,</td>
<td></td>
<td>(ST22, 25)</td>
</tr>
<tr>
<td></td>
<td>RSS290 (ST8, 8, 1,</td>
<td></td>
<td>NRS158 (ST1, 1, 88, 22, 1, 1, 8,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30, 121, 30, 30)</td>
<td></td>
<td>22, 80)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The competitive ELISA protocols were carried out in two batches (Set 1 and Set 2), as described in Section 5.2.6. Isolates producing toxins in each range of concentrations are as indicated with the actual lowest and highest concentrations detected within each range shown above the isolate identity. The concentration values for each range were arbitrarily decided.

The sequence type of each isolate is indicated below the isolate identity.
Table 5.3: Comparison of LukF and HLA protein production for selected isolates

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID</th>
<th>LukF Concentration (µg/ml)</th>
<th>HLA Concentration (µg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS1</td>
<td>0.27</td>
<td>0.12</td>
</tr>
<tr>
<td>2.</td>
<td>TS8</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>3.</td>
<td>TS16</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>4.</td>
<td>TS17</td>
<td>4.34</td>
<td>4.48</td>
</tr>
<tr>
<td>5.</td>
<td>TS19</td>
<td>0.97</td>
<td>2.07</td>
</tr>
<tr>
<td>6.</td>
<td>TS24</td>
<td>7.91</td>
<td>5.06</td>
</tr>
<tr>
<td>7.</td>
<td>TS25</td>
<td>0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>8.</td>
<td>TS28</td>
<td>0.10</td>
<td>0.21</td>
</tr>
<tr>
<td>9.</td>
<td>TS29</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>10.</td>
<td>TS30</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>11.</td>
<td>NRS194</td>
<td>0.05</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Protein concentrations were defined as similar if they had an approximate difference of 2 fold or less

ND: Not Detected i.e. below level of detection of the assay
5.4 Discussion

5.4.1 ELISA protocol

Currently, confirmation of *S. aureus* isolates as PVL-positive is carried out by the HPA SRU, based on the presence of the genes encoding the PVL toxin. This does not however provide any information regarding levels of toxin gene expression. The present study set out to describe an ELISA capable of detecting and quantifying PVL toxin production, with the potential for use as a diagnostic and research tool. At the start of the present study, no PVL-ELISA protocol had been specifically described designed for use as a diagnostic tool. Two publications had however made mention of PVL ELISA (Hamilton et al., 2007, Tseng et al., 2009). In these studies, ELISA had been used as a research tool in quantifying PVL toxin production *in vitro*, as part of a larger study. The ELISA protocol described in the present study differs from these other studies in several respects.

The first variation was in design. The ELISA protocols used in all three studies had a similar number of steps but differed in specific format. A competitive ELISA protocol comprised of 5 steps was used in the present study. In contrast, Hamilton et al. (2007) used a 6 step sandwich ELISA protocol. This protocol additionally required a 5-fold sample concentration prior to the start of the assay and each ELISA step involved a 2 h incubation. Comparatively therefore, the ELISA described in this present study was simpler and quicker in its execution. No sample concentration was required, and incubation times for each step ranged from 0.5 h to 1 h. In addition, in the present study, a lower threshold detection level of 7 ng/ml was obtained in comparison to a
threshold detection limit of 36 ng/ml reported by Hamilton et al. (2007). On the other hand, the ELISA protocol described by Tseng et al. (2009), which involved the use of a 4 step indirect ELISA protocol, showed a better threshold detection limit (2 ng/ml) than noted in the present study.

One other significant difference between the ELISA protocol described in the present study and the earlier reported PVL-ELISA protocols was the type of primary antibody employed. Unlike both the Hamilton and Tseng studies which made use of a polyclonal primary capture antibody, a monoclonal primary capture antibody was used in the present study. Polyclonal antibodies are generated by immunisation of experimental animals with the appropriate antigen. The resulting antiserum, contains a mixture of specific antibodies which are derived from different B lymphocyte clones which recognise a range of epitopes present on the antigen (Nelson et al., 2000). These antibodies offer the advantages of being relatively cheap and easy to generate and high titre sera can be obtained. Two significant drawbacks are however associated with these antibodies. The polyclonal nature of the antibodies present results in a higher potential for cross-reactivity with other antigens present in test samples, and this could reduce assay specificity. Secondly, as no mechanism exists to ensure that the antibodies produced during each round of immunization and in different animals are against identical epitopes (Hjelm et al., 2012), the use of polyclonal antibodies is affected by potential batch to batch variations.

In contrast, monoclonal antibodies are by definition antibodies produced by a single B lymphocyte clone against a specific epitope. These antibodies are classically generated using the hybridoma technology developed in the 1970s (Kim et al., 2012,
Kohler and Milstein, 1975). This method involves the fusion of splenocytes from immunized animals with immortal mouse myeloma cell lines capable of producing antibodies. Following this, hybrid cells are selected and screened for a specific antibody producing clone. These cells can then be grown and the monoclonal antibody harvested from the supernatant. Thus, although initially more technically demanding to produce, use of monoclonal antibodies reduces the risk of cross-reaction, eliminates the possibility of batch to batch variation and ultimately reduces use of experimental animals.

A third PVL-ELISA protocol has also been described. This protocol which was briefly summarised in 2008 (Badiou et al., 2008), was only recently fully described and further evaluated for the detection of the PVL toxin in clinical samples (Badiou et al., 2010). This system is however not yet commercially available. Unlike previous PVL-ELISA protocols, this 4 step sandwich ELISA detects the LukS rather than LukF protein and also uses a monoclonal antibody. It has incubation times for each step in the protocol ranging from 0.5 h to 1.5 h and a 5 ng/ml detection limit.

The competitive ELISA protocol used specifically in the present study involved the use of a biotinylated protein/antigen detected via a streptavidin-HRP conjugate. This assay could thus be further described as ‘a biotin-avidin (BA) ELISA’. This labelled avidin-biotin (LAB) technique which was first described in 1979 (Guesdon et al., 1979), takes advantage of the very strong almost irreversible high affinity binding (Gyorgy and Rose, 1941, Green, 1963, Melamed and Green, 1963, Chen et al., 2012) found to occur between the 244.31 Da molecular weight biotin molecule and the
≈67 kDa egg protein avidin. The very small size of biotin enables the labelling of either antibodies or antigen without any effect on activity of these molecules. The very highly specific avidin-biotin affinity additionally ensures binding to only target molecules, thus reducing non-specific binding and improving specificity (Diamandis and Christopoulos, 1991).

The use of streptavidin rather than avidin is a more recent variation of the avidin-biotin interaction. Streptavidin has similar binding characteristics for biotin as avidin (Chaiet and Wolf, 1964, Tausig and Wolf, 1964) and is produced by *Streptomyces avidinii*. This protein however differs in amino acid sequence from avidin. Unlike avidin, streptavidin has no carbohydrate content resulting in lower non-specific binding when compared with avidin (Diamandis and Christopoulos, 1991, Kurstak, 1985).

Other advantages of the LAB technique include an increase in ELISA sensitivity when compared with a standard ELISA process involving the detection of an immune reaction simply using enzyme labelled antibodies or antigen. This increased sensitivity was demonstrated in a 1983 study designed to detect anti-hepatitis B surface antigen antibody in mice (Kendall et al., 1983). In this study, a nearly 100 fold difference was found between standard and biotin-avidin (BA) ELISAs. The standard ELISA was also 50 times more sensitive than a commercial radioimmunoassay (AUSAB) test while the BA ELISA was 4,134 times more sensitive than this standard. Similarly, other studies reported increased sensitivities of 8-fold (Wang et al., 2010) and 32-fold (Edwin, 1989) when compared with non-BA-ELISA systems.
5.4.2 **Recombinant protein production**

The general need for large quantities of microbial proteins in both therapeutics and research naturally drove the science of recombinant gene cloning, protein expression and protein purification. Several different expression systems exist and these rely on a variety of promoter systems which differ in their level of expression (Terpe, 2006, Baneyx, 1999). Compared to other promoter systems such as the L-arabinose inducible $P_{BAD}$ promoter, the $trc$ and $tac$ promoter, the lambda phage promoter and the anhydrotetracycline-inducible promoter, the T7 phage promoter employed in the pET expression system is known for its very high levels of expression (Terpe, 2006).

The suitability of the pET expression system in the cloning and expression of both LukF and HLA-H35L proteins was noted in the present study. Despite the widespread use of the pET system, it has been found unsuitable in some cases by failing to result in protein expression. This problem usually arises during the production of membrane proteins or proteins lethal to the cell (Baneyx, 1999). One such failure of the pET system was reported during the production of the adenylate cyclase enzyme where no protein was detected following cloning and expression of the enzyme genes (Reddy et al., 1989). This phenomenon was thought in part, to be due to issues of leaky T7 RNA polymerase expression. Leaky or uninduced expression has been associated with promoters derived from the lac operon (Giacalone et al., 2006, Grossman et al., 1998). Leaky expression occurs when the promoter is not tightly regulated and refers to the background expression of an inducible promoter in the absence of its inducer. This phenomenon could lead to plasmid instability with or without a loss of the plasmid, thus negatively impacting on the level of expression. Furthermore, leaky expression
may equally result in inconsistent levels of protein expression in some cases (Paul et al., 1997, Mertens et al., 1995, Spehr et al., 2000).

Methods to either overcome or bypass this limitation have included the use of low levels of phage T7 lysozyme. Phage T7 lysozyme degrades the T7 RNA polymerase, thereby inhibiting its activity (Studier, 1991, Baneyx, 1999). In addition to such modifications to existing systems to create a tighter regulation and prevent leaky expression, the use of alternate promoter systems specifically designed to prevent leaky expression, has been reported. Examples of these include the bacteriophage lambda promoter (PL) (Elvin et al., 1990, Terpe, 2006) the L-rhamnose inducible promoter (Prha) (Haldimann et al., 1998), the L-arabinose inducible promoter (PBAD) (Guzman et al., 1995) and the hybrid lac and arabinose promoter (Plac/ara-1) (Lutz and Bujard, 1997) based systems. In this present study however, none of these issues were observed to pose a problem. This was reflected by a lack of background expression in the uninduced state (which could cause toxicity and hinder recovery of target protein), and the rapid induction of expression with high levels of protein produced within 2 h as depicted in Figure 5.7 and Figure 5.8.

On the other hand however, the significant level of insolubility of the LukF protein following induction at 37°C (Figure 5.8), represented a different set of issues linked with the pET expression system. The problem of formation of insoluble inclusion bodies has long been associated with this system (Paul et al., 1997). This problem ironically is associated with the high level of protein expression achieved by the system (Baneyx, 1999). Inclusion body formation has been linked with overproduction of a protein. As higher growth rate is linked to increased temperatures,
insolubility has been shown to be enhanced when higher growth temperatures were used (Klein and Dhurjati, 1995). Conversely, protein induction at lower temperatures has been shown to result in increased protein solubility, though this is often associated with reduced levels of protein expression. This phenomenon, which has been reported for a variety of proteins (Chalmers et al., 1990, Weickert et al., 1997, Back et al., 1994), was also observed in the present study (Figure 5.8).

Protein purification in the present study was carried out using kits which rely on immobilised metal affinity chromatography (IMAC) (Porath et al., 1975) for the separation of target protein from all other proteins in solution. This method exploits the high affinity of specific amino acids (histidine and cysteine) for transition metal ions such as Zn\(^{2+}\), Cu\(^{2+}\), Ni\(^{2+}\) and Co\(^{2+}\) (Smith et al., 1988). A widely used application of this is the purification of a 6 × his-tag fused to a recombinant protein using nickel charged columns (Block et al., 2009) as used in the present study. His-tags are added as part of the cloning process. For example, the pET-21d(+) plasmid expression vector used in the present study for the cloning and expression of both proteins, is designed to incorporate a C'-terminal hexa-histidine tag onto the protein, thereby enabling ease of purification. The affinity of histidine for nickel is both specific and reversible (Cuatrecasas, 1970). Additionally, as this tag is uncharged at physiological pH and generally poorly immunogenic it has generally been found to have little effect on the properties of the tagged protein (Crowe et al., 1994, Bornhorst and Falke, 2011).

While different commercial purification protocols rely broadly on the principles of IMAC, the specifics of the process may vary from one to another. In the present study, improved purification of the HLA-H35L protein was achieved with the BioRad®
Biologic LP affinity chromatographic system. The significant difference between this system and the Novagen His Bind Resin® protein purification kit is in the use of a linear imidazole gradient for protein elution as opposed to the stepwise elution used by the Novagen kit. The stepwise elution process limits protein elution to a single high concentration of imidazole (250 mM) and works effectively for many proteins, as was observed with the LukF purification (Figure 5.9). Gradient elution however, takes into consideration the fact that all bound proteins do not necessarily have similar elution profiles, and enables contaminating proteins bound non-specifically to be washed out at lower imidazole concentrations. This highlights how slight variations of IMAC protocols may impact on the efficacy of purification of different proteins.

5.4.3 Summary

Overall, this chapter described the development of sensitive ELISA protocols for the detection and quantification of LukF and HLA proteins using monoclonal antibodies. These preliminary results generated following the development and validation of both immunoassays indicate the potential of the assays, both for clinical diagnosis and as research tools to detect and investigate the expression of these toxins. Further validation would however be essential before this potential can be achieved.

One significant variation in the results generated in this chapter from those of previously published studies (Hamilton et al., 2007, Badiou et al., 2010, Loughman et al., 2009, Said-Salim et al., 2005), was the inability to detect any expressed LukF protein in three PCR confirmed PVL-positive strains. This potentially false negative finding could drastically alter the potential application of ELISA in clinical diagnosis of PVL-positive isolates. A comprehensive follow-up study would be needed to
quantitatively explore the effect of various previously reported parameters on toxin expression (Section 1.6.6) to explore the possibility of improving the potential utility of the assay.

For the other isolates tested, the ELISA assays showed clear differences in expression levels of both LukF and HLA. In some cases, similar levels of both proteins were produced by individual isolates but for others, expression levels did not appear to be linked. These data suggest that factors controlling toxin expression vary between different isolates though the mechanisms involved are currently not defined and require additional study.
Chapter Six

6 General Discussion

6.1 Overview

The ability of *S. aureus* to cause infection in a wide variety of tissues and the extensive array of virulence determinants it possesses has been well documented (Archer, 1998, Casey et al., 2007, Lowy, 1998). These characteristics, together with its penchant for development of antibiotic resistance have combined to make this species a significant cause for concern, with the hospital acquired strains proving increasingly problematic for over four decades. Up to 2001, the voluntary MRSA reporting scheme begun in England in the 1990s had showed a steady increase in the number of reported MRSA bacteraemia cases. Over this time period, the number of reported cases had risen to 5,000 per year. Data collected following the introduction of mandatory MRSA reporting in England in 2001 (Department-of-Health, 2001), confirmed the high prevalence of MRSA bacteraemia cases (Figure 6.1), with over 7,000 cases per year reported by 2002. This therefore, led to a government intervention in 2003 to reduce the incidence of MRSA bacteraemia (HPA, 2004), and a target of a 50% reduction in MRSA bacteraemia cases by 2008 was set. Despite the possibility that the rise in number of cases was an anomaly due to increased awareness and surveillance, by 2008, the government target of a 50% reduction in MRSA bacteraemia cases was achieved (Figure 6.1 and Figure 6.2). Furthermore, as the data from the voluntary surveillance showed (Figure 6.2), this reduction was specific for bacteraemia caused by MRSA cases rather than MSSA bacteraemia cases in general. Numbers of bacteraemia cases caused by MSSA showed a slight increase over this
period. Notwithstanding the further steady decline in incidence of MRSA bacteraemia cases reported by the HPA (Johnson et al., 2012), the fight against *S. aureus* is not yet over, as once again the ‘modus operandus’ of this ever versatile species appears to be changing.
Figure 6.1: Incidence of reported MRSA bacteraemia in England (1993 – 2008)

The figure represents data on the number of MRSA bacteraemia reports obtained initially via the voluntary MRSA surveillance scheme only (1993 – 2002), and then also from the with mandatory MRSA surveillance scheme (2002 – 2008). The figure shows a similar trend between voluntary and mandatory surveillance reports from 2002 and a steady post 2003 decline in reported cases.

The figure is reproduced from Pearson et al. (2009).
Figure 6.2: Incidence of reported *S. aureus* bacteraemia in the UK (2002 – 2011).

The graph shows data on the number of cases of bacteraemia caused by MSSA, MRSA and total *S. aureus* from the voluntary reporting scheme. This data shows consistency of reported numbers of MSSA bacteraemia over the reporting period.

The changing face of *S. aureus* was first observed in the 1990s with initial reports of MRSA causing infection in the community (Herold et al., 1998, Price et al., 1998, CDC, 1999, L'Heriteau et al., 1999, Gwynne-Jones and Stott, 1999, Berman et al., 1993). These new “CA-MRSA” strains were then subsequently disseminated worldwide. The epidemiology of infections caused by these *S. aureus* strains has since evolved further. CA-MRSA strains have now spread into healthcare facilities and have been reported as causing nosocomial infections (Saunders et al., 2007, David et al., 2006, Patel et al., 2007, Valsesia et al., 2010). More recently, there has also been an increase in reports of specific MRSA clones colonising or infecting livestock. These livestock associated MRSA (LA-MRSA) isolates have also been reported in humans exposed to various animals (Garcia-Graells et al., 2012, Richter et al., 2012, Kock et al., 2012) and have even been isolated from healthcare facilities (Lozano et al., 2012, Kock et al., 2011). Together with the higher numbers of PVL-MSSA strains noted in the UK (Table 1.3) and continuing high incidence of MSSA bacteraemia despite the reduction in MRSA bacteraemia in recent years (Figure 6.2), these observations highlight the continuing evolution of the interaction between *S. aureus* and humans to facilitate the survival of this species in the human population. It is therefore important for the scientific community to keep up with these changes.

This thesis aimed to increase understanding of the role of the PVL toxin in staphylococcal biology, with particular reference to PVL-MSSA isolates and to evaluate novel methods for typing and detection of toxin production by these isolates.
A detailed characterisation of a collection of PVL-MSSA isolates was carried out as part of this research. Generally, this group of isolates appear to have been overlooked in favour of studies with MRSA isolates, with regards to focused epidemiological and typing studies. The gap in accurate epidemiological data for MSSA in general has increasingly been noted, causing the UK Health Protection Agency in January 2011 to introduce a mandatory surveillance system for MSSA bacteraemia cases (HPA, 2011b, Wilson and Kiernan, 2012).

Four different typing systems, in addition to virulence and resistance gene profiling, were employed in the characterisation of PVL-MSSA isolates in the present study. Over the years, the systems used in the characterisation and typing of *S. aureus* have evolved significantly and at present, PFGE, MLST and *spa* typing are the three most commonly used methods described in the literature. Increasingly, researchers do not apply only a single technique to the typing of *S. aureus* isolates. Rather, typing is starting to involve a combination of at least two of these techniques in a bid to improve discrimination (Shore et al., 2010, Ellington et al., 2010, Vindel et al., 2009) and capture both the long and short term evolutionary processes taking place within this group of isolates. One such study demonstrated a higher discriminatory power of 99.98 when PFGE and *spa* were combined, compared with the discrimination of the individual techniques [PFGE (99.69); MLST (94.82); *spa* (97.31)] (Faria et al., 2008).

PFGE was deemed the gold standard in *S. aureus* typing (Peacock et al., 2002a). However, PFGE is cumbersome and the difficulty of intra-laboratory comparisons mean that fewer studies now employ this technique as a typing tool. Rather, the sequence based MLST and *spa* typing systems employed in the present study are
increasingly becoming the techniques of choice, due to a comparable level of discrimination as PFGE. Applying these techniques to the subset of clinical PVL-MSSA isolates in the present study provided a wealth of information. An even more diverse genetic background was noted among these strains than previously reported in PVL-MRSA isolates (Table 3.5 and Table 3.6), especially when compared with the general CA-MRSA population (Muttaiyah et al., 2010, Rasigade et al., 2010, Tong et al., 2010). Interestingly, despite the limited sample size, the test isolates in the present study, appeared to provide a snapshot of trends previously noted among PVL-positive strains in this locale (Ellington et al., 2009, Grundmann et al., 2010, Monecke et al., 2007b, Otter et al., 2009. These PVL-MSSA isolates represented identical clones (as described by both MLST and spa typing) as those previously found in PVL-MRSA isolates, supporting the hypothesis that the latter strains had emerged from the former simply by the acquisition of the SCCmec element (Figure 3.1). As the number of test isolates in the present study was rather limited, further studies typing a larger subset of isolates would be essential to confirm the hypothesis.

Based on MLST and spa typing schemes alone however, it would not be possible to prove that a specific set of PVL-MSSA strains in a locale gave rise to the PVL-MRSA strains circulating in the same population. Theoretically, a stronger evidence to support this hypothesis would be the isolation of both PVL-MSSA and PVL-MRSA strains which are more or less identical not just in terms of their MLST and spa types but also with respect to the variations in PVL encoding genes, type of PVL phage present, and virulence gene profile. Considering the association sometimes of resistance genes with the SCCmec element (Ito et al., 2001), it might also be expected that the evolving PVL-MRSA strains would be more resistant than their meticillin
susceptible counterparts. This is why the other two typing schemes employed in the present study (PVL sequence and PVL phage typing) were used to clearly demonstrate a possible evolutionary link between isolates in the present study and others as previously described (Ellington et al., 2009, Ellington et al., 2010).

Subsequent to the present study, further research in other geographical locations would however need to be performed to determine if indeed these findings are widespread or are limited to the PVL strains found in the UK. Should this hypothesis be confirmed and found applicable to the various PVL-positive strains found worldwide, this could cause a shift in intervention policies leading to proactive as opposed to reactive measures. Rather than looking at ways to manage outbreaks of PVL-positive strains, future interventions could focus on ways to prevent spread of PVL phages within the S. aureus population or target those clones susceptible to infection by these phages. This may involve more directed research looking at PVL phages, their reservoirs and mode of evolution. This may be a more pressing issue than is currently appreciated as in the last year several reports have described different bacteriophages which carry the genes encoding the PVL toxin (Zhang et al., 2011, Le et al., 2012, Ito et al., 2012). One of these studies (Zhang et al., 2011) described two novel PVL phages in a single clone (ST59) with 99% identity between them. The genomes of these phages demonstrated a high degree of homology to previous PVL phages with respect to five genes located at the end of the prophage (int (integrase), hol (holin), ami (amidase), lukS-PV, and lukF-PV). They are however otherwise quite different from that of previously published PVL phages. These novel PVL phages exhibit a higher degree of homology to non-PVL group 3 Sfi21-like Siphoviridae phages (Zhang et al., 2011). This led the researchers to propose that novel PVL
phages may simply be generated by what they termed ‘illegitimate recombination’ involving the uptake of the region containing the five genes mentioned above.

One issue the present study did highlight was the need for more comprehensive typing of *S. aureus* isolates. Apart from the group of isolates in the present study which had novel *spa* or MLST types described for the first time, there was also another set of isolates from which conclusions could not be drawn simply because no comparable data existed for their PVL-MRSA counterparts, or in some cases, this data was incomplete. An analysis of the literature revealed that while the use of more than one typing scheme is becoming more common, quite a number of studies have still failed to appreciate the need for this approach, particularly the necessity of *spa* typing. In a sense, MLST and PFGE are two ends of the typing spectrum. MLST involves an analysis of 7 housekeeping genes thought to evolve slowly over time (Enright et al., 2000). Hence the information it provides is more global, long term epidemiology. PFGE on the other hand involves an analysis of total DNA and has a discriminatory power such that a single genetic change could result in a change in banding pattern and hence data interpretation (Bannerman et al., 1995). On the typing spectrum then, the information provided by *spa* typing falls between these two methods as reflected in its discriminatory power being higher than that of MLST but lower than PFGE (Faria et al., 2008).

The analysis of a collection of MSSA-PVL-positive isolates found in the UK as carried out in the present study, clearly demonstrates the value of *spa* typing. While one of the predominant clones represented among UK PVL-positive isolates is the same clone to which the notorious EMRSA-15 responsible for hospital outbreaks
belongs (Holmes et al., 2005, Ellington et al., 2009, Boakes et al., 2011b), their associated _spa_ types differ. Whereas t032 and t022 are the more common _spa_ types found within EMRSA-15, ST22 isolates (Aires-de-Sousa et al., 2008, Boakes et al., 2011b, Otter and French, 2008), these _spa_ types are however not commonly represented among the ST22 PVL-positive isolates circulating in the UK. A recent study by Boakes and colleagues involved the characterisation of 47 CC22 PVL-MRSA isolates from the HPA collection (Boakes et al., 2011b). None of these isolates belonged to either t022 or t032 _spa_ types. Rather the majority of them (34/47 72.3%) were _spa_ type t005. Therefore, it is recommended that _spa_ typing is regularly incorporated into the typing scheme of _S. aureus_, rather than being seen as an extra adjunct, in order to provide a balance to the information provided by PFGE and MLST.

### 6.2 Novel typing techniques

Despite the advantages associated with both MLST and _spa_ typing, these sequence based typing systems in themselves are still not perfect. While their associated advantages of being less cumbersome, easy of interpretation and portability (especially when compared with PFGE), cannot be ignored, certain drawbacks still exist. With the need for amplification, sequencing and data interpretation of at least 8 loci per isolate, time and financial costs for these methods are significant. Given an increasing number of strains are now being routinely subjected to genotyping, costs continue to escalate. These limitations prompted the investigation of the alternative typing methods described in the present study.
As an emerging technique, high resolution melt (HRM) analysis has shown a great deal of potential with its remarkable degree of sensitivity and specificity (Erali et al., 2008) and its potential for application as a rapid typing system. The high degree of specificity and sensitivity usually associated with HRM, was confirmed by detection of four variations in the lukSF-PV locus (Figure 4.2). Impressively, a great deal of information could be extrapolated from the data generated. In addition to confirming PVL isotypes as previously defined (O'Hara et al., 2008) mutations in the genes encoding both LukF and LukS subunits correlated with ST allowing strain genotyping (Takano et al., 2008b, Berglund et al., 2008b, Otter et al., 2010, Wolter et al., 2007). Based on HRM analysis of variations in the lukSF-PV locus therefore, this technique may be useful in an outbreak setting as a rapid, non-sequence based method for the typing of isolates as same or different if sufficiently discriminatory.

However, attempts to improve discrimination by applying this technique to identifying variations in the spa gene (Figure 4.4) highlighted a number of issues (especially regarding data interpretation), which have not previously been reported in the literature. In the present study, the ultimate aim of using HRM to differentiate spa types was met resulting in improved discrimination for this technique. An analysis of the spa locus was however only possible following an awareness of the level of temperature variability across the plate (Figure 4.11) and the effect of temperature shift on data interpretation (Figure 4.4). Hence, in contrast to the HRM analysis of the lukSF-PV locus where types were automatically assigned, a high degree of user subjectivity would come into play in the interpretation of spa HRM data.
These added issues in data interpretation combined with its poor predictive ability (Table 4.12), might perhaps point to unsuitability of the spa locus as a candidate for HRM. HRM though, has been applied in genotyping based on size variation (Price et al., 2007). While this application is not as widespread as SNP detection and mutation scanning, it has successfully been applied (Stephens et al., 2006). Therefore, the failure to replicate this experiment, albeit on a different HRM platform, highlights a clear problem which could potentially undermine the advantages of HRM, the issue of portability. For a typing system to be successful, one important requirement is for portability so as to enable accurate comparison of data generated by such a technique from in any laboratory worldwide.

In the case of HRM, the availability of several different commercial platforms and associated software (Herrmann et al., 2007), which differ in technology, dye employed and instrument capabilities, acts as a limiting factor. Hence, despite any associated successes in its use as a rapid typing tool, its widespread application presently would be plagued by issues of portability making the technique perhaps useful simply in regional studies. This might perhaps change with improved technologies but this is currently hard to envisage considering that the success of the technique has created a market for the manufacturers and a single supplier of HRM equipment might make the cost prohibitive. Conversely, as the technique becomes more commonplace, enough assays might be carried out on the various platforms making it possible to create a protocol designed to generate comparable results not withstanding equipment variation.
6.3 Future Work

The present study was based on a small collection of PVL-MSSA isolates and provides information which suggests direct evolutionary links between specific UK PVL-MSSA and PVL-MRSA isolates. Further testing of more isolates is however required to determine if these findings are repeated on a larger scale. This would involve the characterisation of a wider collection of national and international PVL-MSSA and PVL-MRSA strains to enable direct comparison. In addition to contributing to the limited current data on what appears to be the emerging PVL-MSSA isolates, the data generated would form the basis for intervention strategies as previously mentioned.

With the high level of potential demonstrated for the application of HRM as a rapid tool for typing of PVL-positive isolates, future work would also be geared towards improving this potential and harnessing it for maximum benefit. In order to do this, it will be necessary to incorporate a third locus into the assay to improve its discriminatory power. The candidate gene for this process would need to be one that is present in all strains of *S. aureus*, in a highly conserved form but with a few variations which correlate with *S. aureus* lineages. The *hsdS* restriction modification genes as described by Lindsay and colleagues (Waldron and Lindsay, 2006) looks like it might fit this profile. Hence the next step would be to explore the ability of HRM to detect variations in this gene in order to enable genotyping.

One factor which could increase the usefulness of this HRM technique would be the use of a single mastermix to provide the same information gleaned from the three individual reactions. This is why the final step in the development of such a rapid test
would be to multiplex all three HRM reactions into one. Multiplexing in relation to HRM has been carried out (Garritano et al., 2009, Senapin et al., 2010, Curd et al., 2011, Chroma et al., 2011), though it has not been commonly reported. In this case, as the annealing temperatures for the primers are already comparable and the expected melt temperatures variable, multiplexing would simply involve optimisation with respect to concentrations to generate individual different peaks which can then be analysed using the software.

Finally, the ELISAs developed in the present study could provide a useful tool for the detection of PVL and HLA expression in research and diagnostic setting. Where previously Western blots had to be used to give a rough estimate of the level of these toxins produced by various isolates, an exact quantitative level can now be achieved. The first application of the ELISA developed in the present study would be to analyse how toxin levels for a large panel of isolates compare not just with the results obtained by Western blotting but also with mRNA transcript levels of the specific genes encoding the toxins. Various publications have cited these methods for assessing PVL expression but no single study has currently used all these techniques on a single set of isolates.

It must be pertinent to note at this point, that there is an emerging field in science which has the potential to change the landscape of diagnosis, making all the techniques in use today either obsolete or second choice. With rapid advances in technology, concomitant cost reduction and increased speed of sequencing (Hall, 2007, Koser et al., 2012a), whole genome sequencing (WGS) is increasingly becoming a force to be reckoned with. Following sequencing of the first bacterial
genome in 1995 (Fleischmann et al., 1995), over 1000 bacterial genomes have now been sequenced. Unlike most typing systems today which simply draw conclusions of evolutionary patterns and relatedness by the analysis of a single gene (as in spa typing), or at best a set of genes (as in MLST), WGS (as the name implies) provides a snapshot of the whole genome, eliminating the need to focus on a specific target and providing an immense amount of information. WGS could thereby find application in more accurate identification and characterisation of isolates, virulence and susceptibility studies as well as in novel drug target detection. This technique has currently been applied in several outbreak investigations (Eyre et al., 2012, Snitkin et al., 2012, Koser et al., 2012b). The development of next generation sequencing (NGS) systems (Niedringhaus et al., 2011, Quail et al., 2012), including benchtop facilities such as Illumina MiSeq which allows samples to be sequenced and data analysed within one day (Liu et al., 2012), looks set to usher in a new era in diagnostics.

WGS has the ability to transform the face of diagnosis and contribute to improving patient care, but the major limitation of the large scale application technology would be the sheer amount of data it generates and associated complexities. While the information it provides is useful, superseding any current technologies, the question would have to asked at some point – is it necessary and does the benefit justify the cost? Answering this in the future, might lead to a balance between old and new technologies to provide healthcare professional with the ‘perfect’ diagnostic system.
6.4 Conclusions

In summary, the present study has generated useful comprehensive typing data for the neglected PVL-MSSA isolates which contributes to the sparse pool of worldwide information. At a local level, it clearly points at an evolutionary link between PVL-MSSA and PVL-MRSA, particularly the ST22 strains, highlighting a need for more concerted efforts aimed at typing these strains and setting in motion intervention strategies. It also highlights the need for alternatives to the current typing systems given how cumbersome it is to generate complete typing information which could be used in evolutionary analysis for just a single isolate.

In addition, despite the issues encountered with HRM at the spa locus, the present study demonstrates the potential the HRM technique shows as a rapid typing tool for genotyping of PVL-positive isolates. And while the potential for this could be improved by the addition of a third locus providing the same ease and clarity as the lukSF-PV locus, at present it could play quite a useful role as a preliminary analysis strategy for typing isolates as same or different. The need for a detailed comparison of the different HRM platforms was also clearly highlighted in the HRM analysis carried out in the present study.

Finally, the present study has described a specific and sensitive ELISA protocol which not only detects the presence of the actual PVL toxin but also allows quantification of toxin concentrations. With the increase in PVL studies and a continuing need to unravel its role in pathogenicity, being able to quantify this toxin is an important part of PVL research which would form the basis for a variety of further studies.
REFERENCES


311


**Staphylococcus aureus** complicating a pilonidal cyst. *Langenbecks Arch Surg*, 392, 761-5.


Montgomery, C. P. & Daum, R. S. 2009. Transcription of inflammatory genes in the lung after infection with community-associated methicillin-resistant


Prevost, G., Jaulhac, B. & Piemont, Y. 1992. DNA fingerprinting by pulsed-field gel electrophoresis is more effective than ribotyping in distinguishing among


Senapin, S., Molthathong, S., Phiwsaiya, K., Jaengsanong, C. & Chuchird, N. 2010. Application of high resolution melt (HRM) analysis for duplex detection of Macrobrachium rosenbergii nodavirus (MrNV) and extra small virus (XSV) in shrimp. Mol Cell Probes, 24, 291-7.


Spence, R. P. 2007. *Development and Validation of a Novel Multi-Well Format DNA Microarray for Diagnosis and Molecular Typing of Staphylococcus aureus infections*. PhD PhD, University of Nottingham.


Appendices

Appendix 1: Media Composition

**Brain Heart Infusion broth**

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain Infusion Solids</td>
<td>12.5 g</td>
</tr>
<tr>
<td>Beef Heart Infusion Solids</td>
<td>5.0 g</td>
</tr>
<tr>
<td>Proteose peptone</td>
<td>10.0 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>2.0 g</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>5.0 g</td>
</tr>
<tr>
<td>Di-sodium phosphate</td>
<td>2.5 g</td>
</tr>
</tbody>
</table>

**Luria Bertani Broth**

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>10.0 g</td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>5.0 g</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>10.0 g</td>
</tr>
</tbody>
</table>
### Tryptic Soy Broth

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreatic Digest of Casein</td>
<td>17.0 g</td>
</tr>
<tr>
<td>Papaic Digest of Soybean</td>
<td>3.0 g</td>
</tr>
<tr>
<td>Dextrose</td>
<td>2.5 g</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>5.0 g</td>
</tr>
<tr>
<td>Dipotassium Phosphate</td>
<td>2.5 g</td>
</tr>
</tbody>
</table>

### CYGP broth

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casamino Acids</td>
<td>10.0 g</td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>10.0 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>5.0 g</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>5.9 g</td>
</tr>
<tr>
<td>β-Glycerophosphate (1.5M)</td>
<td>40 ml</td>
</tr>
</tbody>
</table>
Appendix 2: Buffer/Chemical Compositions

SDS-PAGE buffer preparations

SDS PAGE resolving gel buffer [RGB] (pH 8.8)

<table>
<thead>
<tr>
<th>Component</th>
<th>per 100 ml of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>36.3 g</td>
</tr>
<tr>
<td>1M HCl</td>
<td>48.0 ml</td>
</tr>
</tbody>
</table>

SDS PAGE stacking gel buffer [SGB] (pH 6.8)

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>6.0 g</td>
</tr>
<tr>
<td>1M HCl</td>
<td>48.0 ml</td>
</tr>
</tbody>
</table>

11.5% SDS PAGE resolving gel

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% (w/v) Acrylamide bis</td>
<td>3.6 ml</td>
</tr>
<tr>
<td>RGB</td>
<td>1.25 ml</td>
</tr>
<tr>
<td>10% (w/v) SDS</td>
<td>100 µl</td>
</tr>
<tr>
<td>10% (w/v) Ammonium persulphate</td>
<td>150 µl</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>4.5 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>7.5 µl</td>
</tr>
</tbody>
</table>
### SDS PAGE stacking gel

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% (w/v) Acrylamide bis</td>
<td>750 µl</td>
</tr>
<tr>
<td>SGB</td>
<td>1.5 ml</td>
</tr>
<tr>
<td>10% (w/v) SDS</td>
<td>60 µl</td>
</tr>
<tr>
<td>10% (w/v) Ammonium persulphate</td>
<td>150 µl</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>3.0 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>6 µl</td>
</tr>
</tbody>
</table>

### SDS PAGE running buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>14.4 g</td>
</tr>
<tr>
<td>Tris</td>
<td>3.03 g</td>
</tr>
<tr>
<td>SDS</td>
<td>1 g</td>
</tr>
</tbody>
</table>

### SDS PAGE destain (Methanol/Acetic Acid)

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanol</td>
<td>200 ml</td>
</tr>
<tr>
<td>Glacial Acetic acid</td>
<td>75 ml</td>
</tr>
</tbody>
</table>
### 2× SDS PAGE sample buffer (pH 6.8)

<table>
<thead>
<tr>
<th>Component</th>
<th>per 100ml of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>1.5 g</td>
</tr>
<tr>
<td>SDS</td>
<td>4 g</td>
</tr>
<tr>
<td>Mercaptoethanol</td>
<td>10 ml</td>
</tr>
<tr>
<td>Sucrose</td>
<td>20 g</td>
</tr>
<tr>
<td>Bromophenol Blue</td>
<td>0.004 g</td>
</tr>
</tbody>
</table>

### 50× TAE Buffer (pH 8.0)

<table>
<thead>
<tr>
<th>Component</th>
<th>per litre of water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>242 g</td>
</tr>
<tr>
<td>Glacial acetic acid</td>
<td>57.1 ml</td>
</tr>
<tr>
<td>0.5M EDTA</td>
<td>100 ml</td>
</tr>
</tbody>
</table>
### Appendix 3: Identity of Blinded Isolates

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TS1</td>
</tr>
<tr>
<td>2.</td>
<td>TS6</td>
</tr>
<tr>
<td>3.</td>
<td>TS13</td>
</tr>
<tr>
<td>4.</td>
<td>TS15</td>
</tr>
<tr>
<td>5.</td>
<td>TS5</td>
</tr>
<tr>
<td>6.</td>
<td>TS16</td>
</tr>
<tr>
<td>7.</td>
<td>TS18</td>
</tr>
<tr>
<td>8.</td>
<td>TS8</td>
</tr>
<tr>
<td>9.</td>
<td>TS9</td>
</tr>
<tr>
<td>10.</td>
<td>TS2</td>
</tr>
<tr>
<td>11.</td>
<td>TS17</td>
</tr>
<tr>
<td>12.</td>
<td>TS12</td>
</tr>
<tr>
<td>13.</td>
<td>TS7</td>
</tr>
<tr>
<td>14.</td>
<td>TS14</td>
</tr>
<tr>
<td>15.</td>
<td>TS20</td>
</tr>
<tr>
<td>16.</td>
<td>TS23</td>
</tr>
<tr>
<td>17.</td>
<td>TS21</td>
</tr>
<tr>
<td>18.</td>
<td>TS24</td>
</tr>
<tr>
<td>19.</td>
<td>TS19</td>
</tr>
<tr>
<td>20.</td>
<td>TS25</td>
</tr>
<tr>
<td>21.</td>
<td>TS26</td>
</tr>
<tr>
<td>22.</td>
<td>TS27</td>
</tr>
<tr>
<td>23.</td>
<td>TS28</td>
</tr>
<tr>
<td>24.</td>
<td>TS29</td>
</tr>
<tr>
<td>25.</td>
<td>TS30</td>
</tr>
<tr>
<td>26.</td>
<td>NRS123</td>
</tr>
<tr>
<td>27.</td>
<td>NRS157</td>
</tr>
<tr>
<td>28.</td>
<td>NRS162</td>
</tr>
<tr>
<td>29.</td>
<td>NRS185</td>
</tr>
<tr>
<td>30.</td>
<td>NRS158</td>
</tr>
<tr>
<td>31.</td>
<td>NRS192</td>
</tr>
<tr>
<td>32.</td>
<td>NRS194</td>
</tr>
<tr>
<td>33.</td>
<td>NRS227</td>
</tr>
<tr>
<td>34.</td>
<td>NRS229</td>
</tr>
<tr>
<td>35.</td>
<td>NRS248</td>
</tr>
<tr>
<td>36.</td>
<td>NRS255</td>
</tr>
<tr>
<td>37.</td>
<td>RSS043</td>
</tr>
<tr>
<td>38.</td>
<td>RSS063</td>
</tr>
<tr>
<td>39.</td>
<td>RSS289</td>
</tr>
<tr>
<td>40.</td>
<td>RSS290</td>
</tr>
</tbody>
</table>
Appendix 4: Demonstration of lack of haemolytic activity of HLA-H35L on blood agar plate

The haemolytic activity of the recombinant HLA-H35L toxin was assessed on horse blood agar. A 20 µl volume of purified, dialysed protein (100 µg/ml) was pipetted into wells cut in the agar plate. Purified *S. aureus* alpha toxin (Sigma-Aldrich, Dorset UK) (100 µg/ml) was used as a positive control. The plate was incubated overnight at 37°C and examined for zones of haemolysis.
Appendix 5: Raw ELISA results

<table>
<thead>
<tr>
<th>S/No</th>
<th>Isolate ID</th>
<th>Concentration (µg/ml)</th>
<th>LukF</th>
<th>HLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TS1</td>
<td>0.27</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>TS2</td>
<td>0.39</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>TS5</td>
<td>0.03</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>TS6</td>
<td>0.60</td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>TS7</td>
<td>0.01</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>TS8</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>TS9</td>
<td>0.15</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>TS12</td>
<td>0.01</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>TS13</td>
<td>0.04</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>TS14</td>
<td>1.11</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>TS15</td>
<td>0.03</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>TS16</td>
<td>0.08</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>TS17</td>
<td>4.34</td>
<td>4.48</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>TS18</td>
<td>1.37</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>TS19</td>
<td>0.97</td>
<td>2.067</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>TS20</td>
<td>1.32</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>TS21</td>
<td>0.07</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>TS23</td>
<td>1.27</td>
<td>4.62</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>TS24</td>
<td>7.91</td>
<td>5.06</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>TS25</td>
<td>0.08</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>TS26</td>
<td>0.15</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>TS27</td>
<td>0.11</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>TS28</td>
<td>0.10</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>TS29</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>TS30</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>NRS123</td>
<td>0.06</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>NRS157</td>
<td>0.66</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>NRS158</td>
<td>0.15</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>NRS162</td>
<td>0.06</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>NRS185</td>
<td>1.58</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>NRS192</td>
<td>0.04</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>NRS194</td>
<td>0.05</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>NRS227</td>
<td>0.61</td>
<td>8.52</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>NRS229</td>
<td>ND</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>NRS248</td>
<td>0.04</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>NRS256</td>
<td>0.06</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>RSS043</td>
<td>ND</td>
<td>6.70</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>RSS063</td>
<td>ND</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>RSS289</td>
<td>0.04</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>RSS290</td>
<td>0.07</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>

ND = Not Detected
Molecular Characterization and Panton-Valentine Leucocidin Typing of Community-Acquired Methicillin-Sensitive Staphylococcus aureus Clinical Isolates

Kome Otokunefor, Tim Sloan, Angela M. Kearns and Richard James

Published Ahead of Print 20 June 2012.

Updated information and services can be found at: http://jcm.asm.org/content/50/9/3069

These include:

REFERENCES
This article cites 30 articles, 18 of which can be accessed free at: http://jcm.asm.org/content/50/9/3069#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to to another ASM Journal go to: http://journals.asm.org/site/subscriptions/

Journals.ASM.org
Molecular Characterization and Panton-Valentine Leucocidin Typing of Community-Acquired Methicillin-Sensitive Staphylococcus aureus Clinical Isolates

Kome Otokunefor,1,2 Tim Sloan,3 Angela M. Kearns,4 and Richard James5
Centre for Healthcare Associated Infections, School of Molecular Medical Sciences, University of Nottingham, Centre for Biomolecular Sciences, University Park, Nottingham, United Kingdom;4 Department of Microbiology, Faculty of Science, University of Port Harcourt, Port Harcourt, Nigeria;5 and Staphylococcus Reference Unit, Microbiology Services Colindale, Health Protection Agency, London, United Kingdom.

Limited comprehensive molecular typing data exist currently for Panton-Valentine leucocidin (PVL)-positive, methicillin-sensitive Staphylococcus aureus (PVL-MRSA) clinical isolates. Characterization of PVL-MRSA isolates by multilocus sequence typing (MLST) and spa typing in this study showed a genetic similarity to PVL-positive, methicillin-resistant S. aureus (PVL-MRSA) strains, although three novel spa types and a novel MLST (ST1518) were detected. Furthermore, the detection of PVL phages and haplotypes in PVL-MRSA identical to those previously found in PVL-MRSA isolates highlights the role these strains may play as precursors of emerging lineages of clinical significance.

The 1999 CDC report of four pediatric deaths involving the so-called community-acquired, methicillin-resistant Staphylococcus aureus (CA-MRSA) isolates and the association of these strains with Panton-Valentine leucocidin (PVL) (8, 18) was followed by the rapid emergence of CA-MRSA, fueling the growing incidence of MRSA worldwide. CA-MRSA strains are striking in their ability to cause infection in young, apparently healthy, immunocompetent hosts, sometimes with severe and fatal outcomes. While the strong epidemiological link between PVL and CA-MRSA is compelling, the precise role of the toxin in virulence and pathogenesis is yet to be elucidated. This bicomponent, pore-forming toxin, encoded by a highly conserved ~1.9-kb lukSF-PV locus consisting of two adjacent, cotranscribed lukF and lukS genes (26), has 12 major single nucleotide polymorphisms (SNPs), the majority of which are synonymous. A nonsynonymous mutation at position 527, however, serves as the basis of the H and R isoforms (2, 12, 22, 28). Currently, though most research is focused on PVL-positive MRSA (PVL-MRSA), a rising incidence of PVL-positive, methicillin-sensitive S. aureus (PVL-MRSA) infections reported recently is the main contributing factor in the increased incidence of PVL-positive strains in some locales, with approximately 60% of total PVL-positive S. aureus isolates in England in the past 5 years found to be susceptible to methicillin (6, 16, 29). With clinical and epidemiological characteristics similar to those of CA-MRSA (9, 25), PVL-MRSA may represent a hitherto-unrecognized, overlooked emerging public health threat. Several studies have recently attempted to address this information imbalance (7, 21, 27, 29); however, only two of these reported on PVL gene polymorphisms and their implications (7, 29). In addition to contributing to the limited molecular typing data on PVL-MRSA strains, this study sought to explore PVL gene polymorphisms and phage distribution in this group and how this relates with those previously observed in PVL-MRSA strains. This would aid current understanding of the evolution and emergence of PVL-positive CA-MRSA isolates and help to more accurately assess the current threat posed by these strains.

Nineteen PVL-MRSA clinical isolates recovered by the microbiology laboratory at the Nottingham University Hospitals NHS Trust (NUH) based on either clinical suspicion or an antibiogram of gentamicin/trimethoprim resistance which had been associated locally with PVL positivity, submitted to the Health Protection Agency’s national Staphylococcus Reference Unit (SRU) for PVL testing, were analyzed in this study. The isolates were representative of a range of sample types, clinical histories, and patient ages (Table 1) and had no known epidemiological links, with the exception of two clusters; T56 and T59 were recovered from different patients on the same hospital ward, while TS18 and TS24 were isolated from different samples of an unrelated patient and TS17 from a relative of this patient. The presence or absence of 13 toxin genes (sea, seb, sec, sed, see, seg, seh, seh, sei, sej, tss, eta, etb, and etd) was determined in this population using multiplex PCRs as described previously (1, 20). Sequence-based multilocus sequence typing (MLST) and spa typing methods were used in genotyping of isolates. All isolates were spa typed using the Ridom GmbH spa website, www.spaerver.ridom.de, following sequencing of the variable X region of the spa gene as previously described (13), and multilocus sequence types were mapped via the spa http://spa.ridom.de/mlst.shtml database. Representative isolates were further characterized by MLST (14) via the S. aureus database, http://saureus.mlst.net/. Following amplification and sequencing of two internal fragments of 764 bp and 535 bp in the lukS and lukF loci, respectively, using the primers LukSF (ATGGTCAAAAAAGACTATTACGTG), LukSR (TCAAATTCCGTCTGTACTCCCTAG), LukFF (TCGAGTAGATTATGACCG) and LukFR (nATTTCCATCTTATAAATTACCTAC); PVL types were determined based on these sequences. Specific PVL-encoding phage types were detected by the use of nine PCRs.
TABLE 1 Clinical characteristics of Panton-Valentine leukocidin-
positive Staphylococcus aureus test isolates.

<table>
<thead>
<tr>
<th>Clinical characteristic/disease pathology</th>
<th>No. of isolates for patient age group (yrs)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–19</td>
<td>19–39</td>
</tr>
<tr>
<td>Skin and soft tissue infections</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Bacteremia</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Colonization</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Emphysemia</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

(4, 19) which detect six PVl-encoding phages (Table 2), as well as uncharacterized PVl phages classified as either isochedal or elongated head types. The 19 PVl-MRSA isolates, comprising both hospital- and community-acquired isolates recovered between November 2008 and May 2009 (Table 2), were predominantly from skin and soft tissue infections, with a single isolate (TS1) recovered from a fatal case of necrotizing pneumonia. A low prevalence of virulence/toxin genes was noted in this population, with 10% of all isolates carrying 89.5% (17/19) of isolates. Only sequence type 722 (ST722) isolates encoded up to 4 of the 13 toxin genes analyzed (Table 2); the sel and nx genes were not detected in this study group. In a similar trend, 89.5% (17/19) of isolates were trimethoprim resistant, with additional gentamicin resistance observed in 27.9% (11/19) of isolates. The highest resistance to 4 antibiotics was observed in a single ST19 isolates (TS1). However, all isolates were susceptible to ciprofloxacin, rifampin, lincozamycin, fusidic acid, and teicoplanin. Genotyping using the spa technique revealed 11 types clustered into 2 spa clonal clusters (CCs) (CC005 and CC345/657) and 5 singletons based on the BURP algorithm (StaphType software program; Ridom GmbH, Warzberg, Germany). Most frequent (47.4%) were CC005 and CC021, while 7 other types were represented by only a single test isolate. We identified 3 spa types (t6642, t6643, and t6769) that have not previously been described. Strain diversity was further noted with the detection of 6 MLST STs grouped by eBURST software analysis into 5 CCs of known MRSA lineages (CC1, CC22, CC30, CC88, and CC152). ST722, which has been specifically associated with gentamicin and trimethoprim resistance (5), occurred most frequently (n = 9 [47.4%]). ST1518 (CC152) was identified for the first time ever in this study. This strain, which was isolated from a fatal case of necrotizing pneumonia, is a single-locus variant of ST152 differing by a single mutation in the plc allele.

Compared with the proposed lukSF-PV progenitor (30), a total of seven SNPs were noted in the PVl gene sequences of the study strains, five of which occurred in the lukS locus (Table 2). Most isolates (n = 17; 89.5%) were of the B1 variant as defined by O’Hara et al. (22), with both H1 and H2 groups present. The nonsynonymous nucleotide 527 A-G to T-G mutation which defines the B variant occurred only in a single isolate (TS1). In general, the resulting PVl SNP profiles were MLST specific but not exclusively, ST30 isolates exhibited two different PVl SNP profiles (H1 and H2). The outlier H1 profile carried by TS12 was homologous to that of ST22 and ST88 and was found in 63.2% of isolates (12/19). Four known phage types (ΦPVl, Φ108PVL, ΦSa2 USA, and ΦSa2m) were detected in the strains studied. For both ST722 isolates, all phage PCRs were negative, pointing to an unknown phage type. The phage present in TS12 could only be described based on its morphology. A direct relationship was noted between the phage types and clonal lineage in the majority of cases. While all ST22 isolates carried ΦPVl with its corresponding PVl SNP H2 profile, the ST30 isolates showed more variability, with both ΦPVl and Φ108PVL detected within this group. A general lack of correlation of PVl phage type with PVl SNP profile occurred in this group and in the ST1 and ST88 isolates. Carriage of multiple phage types by a single isolate was noted in TS25. Comparing the epidemiologically related isolates, the ST69 cluster had identical characteristics, and though TS18 and TS24 from the same patient were identical, TS17 from a relative differed at the spa locus (Table 2).

While the molecular epidemiology of CA-MRSA has been explored extensively (10, 11) and interest in CA-MRSA is on the increase, there exists a lack of detailed knowledge on PVl-MRSA, particularly in relation to the genetics of the lukSF-PV locus. Represented among the MSSA isolates in this study were major sequence types (ST22, ST88, ST30, and ST1) associated with PVl-MRSA clones (21, 27, 29). Though the majority of PVl-MRSA clones circulating in the United Kingdom belong to either ST8, ST30, or ST80 (13, 17, 23), the predominant group described in this study, ST22, also constitute a significant burden of infection (13). These strains are genetically similar in susceptibility and toxin profile, as well as spa type, to recently described ST22 PVl-MRSA United Kingdom strains, rather than the predominant hospital-acquired MRSA (HA-MRSA) ST22 EMRSA-15 clone (5). This study observed the previously noted geographical and clonal bias in the distribution of the PVl haplotypes, with the R haplotype limited to a few lineages, predominantly in the United States and more recently Australia (22, 29). Due to previous reports of R haplotypes being found in the United Kingdom in significant numbers of MRSA strains (4, 24), the dearth of R haplotypes in this small study population may reflect more on the mecthinil susceptibility of the isolates rather than geographical location. There is no significant functional or structural differences are thought to result from the arginine (R)-to-histidine (H) mutation present in this haplotype (2, 3); however, is of interest that the only described R haplotype in this study was linked with a fatal case of necrotizing pneumonia in a healthy adult. One evolutionary pathway postulated for the development of CA-MRSA is the acquisition of PVl genes prior to that of the SCCmeC element. This view of a high genetic similarity between PVl-MRSA and MRSA strains (27) was reinforced in this study, where a number of the PVl-MRSA clinical isolates were found to possess spa and toxin gene profiles identical to those of previously described PVl-MRSA clones. The number of MSSA isolates in this study having PVl isoforms identical to those of their previously described MRSA counterparts also lends further credence to this hypothesis (2, 4, 12). This may not be limited to a single evolutionary event. The detection of several PVl phage types observed in the ST30 lineage in this study and described by Boxes and colleagues in two separate studies (4, 5) could perhaps hint at the possibility of multiple phage acquisitions and hence evolutionary events, though this remains to be verified.

In conclusion, due to the gentamicin/trimethoprim resistance bias in isolate selection, this study provides one of the most...
pensive in-depth analyses of PVL-MRSA ST22 isolates incorporating data on the PVL genes to date. It also highlights the important role PVL-MRSA strains generally play as reservoirs for PVL-MRSA strains due to their direct evolutionary links as seen from the typing results. The data emphasize the need for increased surveillance, which may form the basis for intervention strategies and help curb the emergence and clonal expansion of PVL-MRSA.

ACKNOWLEDGMENTS

Kerne Otusonfera is a Commonwealth Scholar, funded by the Department for International Development.

All authors declare that they have no potential conflict of interest.

REFERENCES


