http://researchspace.auckland.ac.nz

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.
Staphylococcus aureus population genetics and immune responses: correlation with ethnic variation in the incidence of bacteraemia

Stephen Robert Ritchie

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Molecular Medicine and Pathology, The University of Auckland, 2010.
I dedicate this work to my late friend Professor Peter Black. Peter always encouraged me to pursue an interest in research and was always excited to hear about progress and offer advice. I miss the enlightening conversations we had in the hallways around the hospital and the medical school.
Acknowledgement

I was fortunate to have the benefit of three supervisors during this PhD study.

Professor John Fraser provided oversight of the study of the sera and his kind and thoughtful encouragement and insights always helped rapidly resolve problems that arose.

Professor Paul Rainey supervised the analysis of the population genetic data, and his enthusiasm, intelligence and patience was required to help me up the huge learning curve required for this aspect of the study.

Associate-Professor Mark Thomas provided general oversight of the PhD programme and mentored me throughout the process. He listened to me and solved any problems that arose and managed to maintain my focus when things didn’t go to plan. It is a great privilege to be able to work with Mark and to continue to benefit from his education, advice and friendship.

A number of other people helped me with this project. Drs David Holland, Selwyn Lang and Susan Taylor helped me to access patients and specimens at Middlemore Hospital. At LabPLUS, Dr Sally Roberts, Maree Gillies, Mary Bilkey and Tracy Bathgate helped me obtain bacterial strains and with phenotypic testing; and Pat Lonsdale helped me to obtain serum samples. Dr Arthur Morris provided data from Diagnostic-Medlab.

Dr Craig Thornley and Dr Ron King [from the Auckland Regional Public Health Service] provided advice regarding data collection and analysis. Dr Alistair Stewart provided statistical advice and Professor Marti Anderson, provided advice regarding the comparison of bacterial populations. Ms Nicola Best helped with collection of nasal swab specimens during a summer studentship. Dr Graeme Lindsay proof read the thesis and provided impeccable advice.

James Alesana-Slater and his supervisor Professor Pauline Norris collaborated with the collection of SA strains in Samoa; James recruited the patients and performed the swabs in Samoa. Helen Heffernen at ESR performed spa typing and pulsed-field gel electrophoresis on selected isolates. Mary Fakahau sent me SA isolates from Tonga and Eka Buadromo helped me to arrange ethical approval to obtain isolates in Fiji, she also provided me with access to her lab in Suva to package the isolates to be sent back to Auckland.

In John Fraser’s lab, Dr Fiona Radcliffe, Dr Ries Langley, Dr Amanda Taylor, Dr Nicola Jackson and Dr Matthew Chung all spent time teaching me mysterious laboratory methods. Fiona Clow managed project funding and tirelessly answered my numerous questions.

In Paul Rainey's lab, Dr Xue-Xian Zhang helped me to trouble shoot issues with PCR and Dr Eric Libby provided advice and explanation regarding mathematical aspects of population genetics.
My colleagues in the Infectious Disease department at Auckland City Hospital tolerated my absence from the roster and helped out with my clinic patients when I wasn’t around.

This work was supported financially by the Health Research Council of New Zealand and the University of Auckland. I owe gratitude to the many cases of staphylococcal disease and their families who were willing to meet with me and talk to me, despite the horrendous illness they were experiencing at the time.

I owe special thanks and gratitude to my family, whose unwavering love and support was provided unconditionally. My children, Ben, Rosie and Ed never complained when I was grumpy or frequently absent; all of you constantly amaze me and I promise to make it up to you! My wife provided support and helped in too many ways to list; Katie, I love you and thank you for allowing me to indulge my aspirations.
Abstract

**Introduction:** *Staphylococcus aureus* [SA] is a common cause of disease in New Zealand, in particular the incidence of SA bacteraemia [SAB] is high. SAB is more common in Māori and Pacific people than people of other ethnicities, but the reasons for this discrepancy are not known.

**Aim:** This thesis investigated a range of reasons for ethnic variation in the incidence of SA infections in Auckland. In particular, the genetic structure of SA populations in Auckland, Samoa, Fiji and Tonga were determined and the adaptive immune responses to SA toxins were determined.

**Materials and Methods:** 150 cases of SAB were interviewed and provided serum samples for analysis. 424 healthy population members were interviewed and cultures were performed of nasal swab specimens. Multi-locus sequence typing was performed on SA isolates from 150 cases of SAB in Auckland, 94 healthy nasal carriers in Auckland, 96 people with skin and soft tissue infection in Samoa, 109 people with predominantly skin and soft tissue infection in Fiji and 18 people with a variety of diseases in Tonga. Radioimmunoassay was used to measure antibody concentrations against toxic shock syndrome toxin [TSST-1] and four staphylococcal superantigen-like proteins [SSL] in acute and convalescent sera from 148 cases of SAB and sera from 21 healthy population members.

**Results:** The incidence of SAB was higher in Māori and Pacific people, but was not caused by a higher prevalence of nasal colonisation or infection with different, more virulent strains of SA. The SA populations in Auckland, Samoa and Fiji were not substantially different; furthermore, the SA populations in the Pacific were not substantially different to SA populations in other parts of the world. Indicators of socioeconomic deprivation were more common in cases of SAB, but Māori and Pacific cases of SAB reported similar access to healthcare as cases of other ethnicities. Sera from Māori and Pacific SAB cases had lower concentrations of neutralising antibody against TSST-1 and SSL9 compared with sera from SAB cases of other ethnicities.

**Conclusion:** Ethnic variation in the incidence of SAB is multi-factorial; socioeconomic deprivation, increased rates of illness requiring invasive medical devices and adaptive immune responses all contribute.
Contents

Acknowledgement.......................................................................................................................... iii

Abstract ........................................................................................................................................ v

Contents ......................................................................................................................................... vi

FOREWORD ..................................................................................................................................... 1

PART A: CLINICAL STUDIES INVESTIGATING THE SPECTRUM OF SA DISEASE IN
AUCKLAND AND THE PACIFIC........................................................................................................ 4

Chapter 1: Ethnic variation in the incidence of Staphylococcus aureus bacteraemia in
Auckland........................................................................................................................................... 4
  1.1 Introduction................................................................................................................................ 4
  1.2 Materials and Methods.............................................................................................................. 6
    1.2.1 Selection of cases of SAB.................................................................................................. 6
    1.2.2 Clinical review.................................................................................................................. 6
    1.2.3 Methods of analysis and definitions used........................................................................... 7
    1.2.4 Statistical methods............................................................................................................ 9
  1.3 Results ....................................................................................................................................... 11
    1.3.1 Age, gender and ethnicity of cases of SAB ..................................................................... 11
    1.3.2 The education level, personal income and occupation of cases of SAB ......................... 14
    1.3.3 The deprivation index of the neighbourhood of domicile of SAB cases ......................... 15
    1.3.4 Alcohol usage and cigarette smoking............................................................................... 16
    1.3.5 The household characteristics of cases of SAB ................................................................. 16
    1.3.6 Access to healthcare......................................................................................................... 17
    1.3.7 Past history of infection caused by SA and comorbid illness......................................... 17
    1.3.8 Past history of infection caused by SA ............................................................................. 18
    1.3.9 Acquisition of SA infection ............................................................................................... 18
    1.3.10 The cause of SAB: diagnosis and MRSA ...................................................................... 19
    1.3.11 Disease severity and length of hospital stay ................................................................. 20
    1.3.12 The outcome and mortality of SAB ............................................................................... 20
    1.3.13 Comparison of indicators of deprivation with population values................................ 22
    1.3.14 Summary of ethnic variation among cases of SAB ......................................................... 22
    1.3.15 The difference between ethnic groups after stratification by age.................................. 24
    1.3.16 The incidence of SAB adjusted for the presence of renal disease and diabetes ........... 24
    1.3.17 Factors associated with community-onset SAB ............................................................ 25
  1.4 Discussion................................................................................................................................... 26

Chapter 2: The burden of S. aureus skin and soft tissue infection in Auckland and
evidence of ethnic variation between subtypes of methicillin-resistant S. aureus causing
wound infection in Auckland.......................................................................................................... 29
  2.1 Introduction................................................................................................................................ 29
  2.2 Materials and Methods.............................................................................................................. 31
  2.3 Results ....................................................................................................................................... 32
    2.3.1 The prevalence and susceptibility of MRSA in Auckland ............................................... 32
    2.3.2 MRSA subtype and age .................................................................................................... 32
    2.3.3 The difference in MRSA subtype by ethnic group ............................................................ 34
    2.3.4 The incidence of nm-MRSA and mr-MRSA infection ...................................................... 35
  2.4 Discussion................................................................................................................................... 37
Chapter 3: The prevalence and diversity of MRSA causing wound infection in Samoa

3.1 Introduction ........................................................................................................... 39
3.2 Aims .................................................................................................................... 40
3.3 Materials and Methods ...................................................................................... 41
  3.3.1 Study procedure .............................................................................................. 41
  3.3.2 Microbiological methods ................................................................................. 41
3.4 Results .................................................................................................................. 43
3.5 Discussion ............................................................................................................. 47

Chapter 4: The prevalence and demographic features of S. aureus carriage in Auckland

4.1 Introduction .......................................................................................................... 49
  4.1.1 Immune and genetic differences between persistent carriers and persistent non-carriers ........................................................................................................... 50
  4.1.2 Bacterial determinants of nasal colonisation ..................................................... 50
  4.1.3 Demographic variation in nasal colonisation ..................................................... 51
4.2 Materials and Methods ....................................................................................... 52
  4.2.1 Study participants ............................................................................................ 52
  4.2.2 Culture of SA .................................................................................................. 52
  4.2.3 Analysis ........................................................................................................... 53
4.3 Results .................................................................................................................. 54
  4.3.1 SA nasal carriage ............................................................................................. 55
  4.3.2 Deprivation and household crowding ............................................................... 55
  4.3.3 Recent healthcare exposure ............................................................................. 56
4.4 Discussion ............................................................................................................. 57

PART B: POPULATION GENETICS OF STAPHYLOCOCCUS AUREUS

Chapter 5: The genetic structure of Staphylococcus aureus populations from the Pacific and comparison with the rest of the world

5.1 Introduction .......................................................................................................... 59
  5.1.1 Multi-locus sequence typing [MLST]: comparison of invasive disease isolates and nasal carriage isolates in Oxford ......................................................... 60
  5.1.2 Differences in invasive SA isolates and nasal carriage isolates of SA .............. 61
  5.1.3 The global population structure of SA ............................................................. 61
  5.1.4 Population genetic study of S. aureus .............................................................. 63
  5.1.5 Mutation, random genetic drift and phylogenies of SA ..................................... 63
  5.1.6 Recombination ............................................................................................... 67
  5.1.7 Selection pressure ........................................................................................... 68
  5.1.8 Migration ......................................................................................................... 69
  5.1.9 Estimating the adequacy of the sample ............................................................ 72
5.2 Aims ...................................................................................................................... 74
5.3 Materials and Methods ....................................................................................... 75
  5.3.1 Identification and testing of SA isolated from nasal swabs ............................... 75
  5.3.2 Selection of SA colonies for genotypic testing .................................................. 77
  5.3.3. Genomic DNA preparation ........................................................................... 77
  5.3.4 Polymerase chain reaction [PCR] and PCR primer design .............................. 77
  5.3.5 Sequence manipulation ................................................................................... 80
  5.3.6 Novel SA strains ............................................................................................. 80
  5.3.7 Software and parameter settings ..................................................................... 81
  5.3.8 The global dataset of SA MLST sequences ...................................................... 84
5.4 Results .................................................................................................................. 86
  5.4.1 The population structure of S. aureus that caused bacteraemia and asymptomatic nasal colonisation in Auckland during 2007 ................................................. 86
    5.4.1.1 The strains of SA isolated from people with bacteraemia ............................ 86
    5.4.1.2 The strains of SA isolated from people with asymptomatic nasal colonisation ..... 86
    5.4.1.3 Polymorphism within the Auckland sample of SA .................................... 87
    5.4.1.4 Identification of a genetic outgroup of SA that caused nasal colonisation .... 88
    5.4.1.5 The Auckland sample of SA is a comprehensive representation of the total SA population in Auckland .......................................................... 89

vii
5.4.1.6 The sample of SA in Auckland would not have improved with additional sampling .......................................................... 90
5.4.1.7 The rates of recombination and point mutation of SA isolated from people living in Auckland .................................................................................. 90
5.4.1.8 Recombination was detected in nucleotide sequences of housekeeping genes of SA ........................................................................................................................................ 90
5.4.1.9 Recombination in the nucleotide sequence alignment of SA has occurred between housekeeping genes ................................................................................................................................. 90
5.4.1.10 SA has lower rates of recombination than Streptococcus pneumoniae and Neisseria meningitidis................................................................................................................. 92
5.4.1.11 The sequence types of SA in Auckland .................................................................................................................................................................................. 92
5.4.1.12 The Auckland SA population lacks genetic diversity compared with outlier strains ........................................................................................................................................ 95
5.4.1.13 The ancestry of the Auckland SA sample: analysis of ST alleles ........................................................................................................ 96
5.4.1.14 The ancestry of the Auckland SA sample: phylogeny derived from nucleotide sequences .......................................................................................................................... 97
5.4.1.15 The ancestry of the Auckland SA sample: the ancestral source of nucleotide polymorphism ........................................................................................................................ 99
5.4.1.16 Evidence of SA evolution in New Zealand ................................................................................................................................. 101
5.4.1.17 The influence of selection on housekeeping genes of SA................................................................................................................................. 103
5.4.1.18 Migration: evidence of transmission of novel sequence types ....................................................................................................... 103
5.4.1.19 Comparison of the sequence types that cause nasal colonisation and bloodstream infection ........................................................................................................ 103
5.4.1.20 Testing the null hypothesis of no difference between the blood and nasal groups of SA .................................................................................................................. 104
5.4.1.21 SA lineage that causes infection in Māori and Pacific people in Auckland .................................................................................. 106
5.4.1.22 Further resolution of STs within CC30 by PFGE did not provide additional correlation with ethnic group or methicillin resistance ............................................................................................................................ 108
5.4.1.23 The analysis of genes in addition to the conventional housekeeping genes does not add to the ability of MLST to differentiate strains of SA ........................................................................................................ 110
5.4.1.24 The functional importance of ssl7 codons under selection pressure ........................................................................................................... 110
5.4.1.25 Co-evolution of ssl7 with SA housekeeping genes ................................................................................................................................. 111
5.4.2 Comparison of the genetic structure of SA populations in Auckland, Samoa, Fiji and Tonga ........................................................ 113
5.4.2.1 The sample of SA strains from Samoa ......................................................................................................................................................... 113
5.4.2.2 The sample of SA strains from Fiji .......................................................................................................................................................... 113
5.4.2.3 The sample of SA strains from Tonga ......................................................................................................................................................... 114
5.4.2.4 The SA samples from Samoa and Fiji were representative of the total SA population in each country .................................................................................................................. 114
5.4.2.5 The STs of SA that cause infection in Samoa and Fiji were similar ........................................................................................................ 116
5.4.2.6 Differences in abundant STs that cause infection in Fiji, Samoa or Auckland ............................................................................... 117
5.4.2.7 The Phylogeny of SA in the Pacific ......................................................................................................................................................... 120
5.4.2.8 Limited evidence of geographical clustering within the phylogeny of SA in the Pacific ......................................................................................................................................................... 120
5.4.2.9 ST93, ST96 and CC121 cause infection only in Māori and/or Pacific people .......................................................................................... 123
5.4.2.10 Minimal difference in the population structure of SA that infect people living in Pacific Island nations or that infect people living in Auckland ........................................................................................................ 124
5.4.2.11 No difference in the diversity of the SA population in different Pacific nations .......................................................................................... 125
5.4.2.12 Statistically significant differences between the genetic structure of SA populations from different Pacific nations ......................................................................................................................................................................................... 126
5.4.2.13 Significant differences in the genetic structure of SA populations from different Pacific nations were caused by minimal differences between populations ......................................................................................................................................................................................... 126
5.4.2.14 Novel SA strains in Pacific Island nations with small human populations ............................................................................................................. 128
5.4.2.15 Migration of novel STs in the Pacific ......................................................................................................................................................... 130
5.4.3 Migration and evolution of SA in the global SA population ................................................................................................................. 130
5.4.3.1 Geographical variation and evolution in SA phylogeny ......................................................................................................................................................... 130
5.4.3.2 Differences between SA populations from Pacific island nations and Mali .................................................................................. 132
5.4.3.3 Geographically diverse SA populations are similar ......................................................................................................................................................... 132
5.4.3.4 No difference between the SA population in Pacific countries and the SA population in Mali ......................................................................................................................................................................................... 134
Chapter 7: Conclusion and future directions

7.1 The prevalence of nasal colonisation might be higher in Māori and Pacific people.
7.2 Colonisation with particularly virulent strains of SA might be more frequent in Māori and Pacific people.
7.3 The use of invasive devices, which increase the risk of invasion by colonising organisms, might be higher in Māori and Pacific people.
7.4 Innate or adaptive immune responses against SA might be less protective against invasive disease.......................................................................................................................... 209
7.5 Māori and Pacific people might develop mild SA disease that might be neglected or undertreated and progress to SAB .......................................................................................................................... 210
7.6 Ethnic variation in rates of SAB might be caused by increased susceptibility and exposure to SA due to socioeconomic deprivation.......................................................................................................................... 210
7.7 Questions for future research .......................................................................................................................... 211

Appendix A: Example of population genetic analysis using ssI11 gene fragment ........ 215
A.1 Example of calculation of the genetic distance between nucleotide sequences ........... 215
A.2 Example rarefaction curve.......................................................................................................................... 216
A.4 Example of Efron’s calculation and Simpson’s index of diversity .................................................. 217
A.5 Example of estimation of recombination and mutation: LDHAT ................................................. 218
A.6 Example of detecting recombination ........................................................................................................ 219
A.7 Example of finding recombination breakpoints using GARD ......................................................... 219
A.8 Example of estimating selection pressure for nucleotide sequences ............................................. 220
A.9 Example of the parsimony test and LibShuff ......................................................................................... 221
A.10 Example of 2 dimensional multi-dimensional scaling ........................................................................ 223
A.11 Example of ANOSIM and PERMANOVA ......................................................................................... 224
A.12 Example of phylogeny reconstruction using maximum likelihood: TREEPUZZLE .... 226
A.13 Example of ancestral population using STRUCTURE ........................................................................ 227

Appendix B: Forms used for data collection ................................................................................................. 229
B.1 Consent form .............................................................................................................................................. 229
B.2 Participant information sheet ................................................................................................................... 231
B.3 Data collection for cases of SAB .............................................................................................................. 235
B.4 Revised consent form for carriage study................................................................................................. 237
B.5 Revised participant information sheet for carriage study ................................................................. 238
B.6 Carriage study data collection form ...................................................................................................... 240

References cited ............................................................................................................................................... 241
Invasive disease caused by *Staphylococcus aureus* [SA] is an alarming illness. Invasive disease is common and can occur in previously healthy people of all ages and from all walks of life. SA infection can be rapidly fatal, and causes serious morbidity and disability in many of those who survive. In many circumstances, such as infection related to a haemodialysis catheter, the cause of disease is obvious; yet in a number of cases there is no simple answer to the question: “Why did this happen to me?”

SA is one of the most important bacterial threats to New Zealanders; in the late 1990s the mortality rate of SA bacteraemia [SAB] was 19%. Cases of SAB outnumbered cases of disease caused by *Neisseria meningitidis* during the epidemic of the 1990s, and the mortality rate of meningococcal disease during the epidemic was approximately 4%.

Of even greater concern was the finding that SAB was more common in Māori and Pacific people living in NZ. There is no immediate explanation for the ethnic variation in SAB incidence although a number of possibilities can be proposed:

1. The ethnic variation in SAB might be the result of confounding caused by other diseases or risk factors. For example, diabetes mellitus is a well known risk factor for SA disease, and it is known that the incidence of type II diabetes varies between ethnic groups. In this thesis, the clinical details of 150 cases of SAB will be examined to determine if other illnesses are the major cause for the ethnic variation in the incidence of SAB.

2. Social factors are likely to play an important part. In young children, household crowding and parental cigarette smoking were important risk factors for the development of meningococcal disease. The social and demographic factors of 150 cases of SAB will be compared with population estimates to examine the importance of deprivation as a risk factor for SA disease. Markers of deprivation will also be compared between SA carriers and non-carriers.

3. Asymptomatic nasal carriage of SA might vary between ethnic groups. Some diseases caused by SA are more common in SA carriers than in non-carriers. Variation in the prevalence of nasal carriage between different ethnicities, might explain variations in invasive disease rates. A large sample from the Auckland population will be examined to determine whether SA carriage is more common in different ethnic groups.

4. Different ethnic groups might be colonised with and have disease caused by different types of SA that vary significantly with regard to their virulence. This might seem an
unlikely proposition, until one considers that infection with Western Samoan phage pattern methicillin-resistant SA [WSPP MRSA] is more common in Māori and Pacific people in Auckland. Laboratory data will be examined to estimate the burden of SA disease in the community and to explore demographic variation in the incidence of MRSA infection. The prevalence and characteristics of MRSA that cause infection in Samoa will be examined to determine whether the prevalence of WSPP MRSA is similar to that in Auckland.

Although many Pacific people living in Auckland maintain strong ties to their country of origin; it is likely that the population of SA varies from one country to the next. The strains of SA that cause infection in Auckland and the Pacific will be examined, at a genotypic level, to determine whether there are important differences between the types of SA causing colonisation or disease in people of different ethnicities in Auckland and three Pacific Island nations.

5. There might be differences in the immune responses to SA colonisation and disease between different ethnic groups. Again, this seems an unlikely proposition, until one considers that the many SA virulence factors which have high degree of specificity for interaction with human cells provides strong evidence that SA has co-evolved with humans over a long period. Thus SAB is likely to have placed a selective pressure on the human population for many generations. This selective pressure may have differed between European and Māori and Pacific people over many generations due to infection by SA strains with major differences in their virulence factors. Immune responses to SA toxins, in patients with SA disease, will be examined to determine whether there are important differences in these immune responses between people of different ethnicities.

Thus, it is the aim of this thesis to explore the magnitude of, and reasons for, ethnic variation in SAB in Auckland. This thesis is presented in three parts that study clinical aspects of SA infection in Auckland and Samoa, the SA populations in Auckland and some Pacific Island nations, and host immune responses to SAB. Each Chapter contains an introduction and discussion; Chapter 7 gives a short conclusion and includes ideas for future research in this field, with the ultimate aim to reduce the incidence of SAB in New Zealand.

Part A: Four clinical studies investigating the spectrum of SA disease in Auckland and the Pacific
Chapter 1: The first clinical study presents a detailed analysis of the demographic and clinical features of 150 cases of SAB admitted to Auckland City Hospital and Middlemore Hospital during 2007. Serum samples collected from these cases were extensively investigated as discussed in Part C; and the SA isolates from these cases were genotyped and analysed in Part B.
Chapter 2: The second clinical study reviewed wound swab data from Auckland’s only community laboratory during 2007, to estimate the burden of staphylococcal skin and soft tissue infection in Auckland and to determine the relative incidence of infection with WSPP MRSA in Māori and Pacific people living in Auckland.

Chapter 3: The third clinical study was performed with investigators at the University of Otago to measure the prevalence and genotypic characteristics of skin and soft tissue infection caused by methicillin-resistant SA in Samoa between November 2007 and February 2008. The SA isolates obtained during this study were used to investigate the population of SA causing disease in the Pacific region [Part B, Chapter 5].

Chapter 4: The fourth clinical study compared the demographic characteristics of SA carriers and non-carriers amongst 424 healthy population members in Auckland during 2007. The genetic structure of the SA populations obtained from this study will be examined in Part B, Chapter 5.

Part B: The genetic structure of SA populations in Auckland and the Pacific region.
Chapter 5: This Chapter presents a detailed analysis of the strains of SA causing colonisation and disease in a number of environments. Multi-locus sequence typing [MLST] was performed on a library of SA collected from asymptomatic nasal carriers and cases of SAB in Auckland; people with wound infections in Samoa; and people with a range of infections in Fiji and Tonga. The genetic structure of the Auckland sample of SA has been determined and the nasal isolates and the bacteraemia isolates compared. The SA sample from Auckland was then compared with the SA sample from Pacific Island nations. The total library of SA from the Pacific region was then compared with reported results from studies employing similar methods to those described in Part A, but conducted in Mali, China and England.

Part C: The immune response to staphylococcal toxins among cases of SAB
Chapter 6: While many studies have investigated human immune responses to SA colonisation and disease, none have shown a consistent correlation with susceptibility to disease [except for toxic shock syndrome toxin, TSST-1]. There is no standardised method to measure the immune responses to SA infection, and there are an enormous number of potential antigens to study. In this Chapter, immune responses to TSST-1 and staphylococcal superantigen-like proteins will be analysed in relation to clinical and demographic features. This study includes serum samples from 148 cases of SAB and 21 healthy population volunteers.

I hope that the data and analysis presented here will make a useful contribution to understanding SA disease in New Zealand and in the Pacific region. I also hope this thesis will assist future research into ethnic variation in a number of other infectious diseases important to New Zealanders. It is unlikely that morbidity and mortality caused by SA has changed dramatically in recent decades, and it is only through research efforts that this situation may change.
Part A: Clinical studies investigating the spectrum of SA disease in Auckland and the Pacific

Chapter 1: Ethnic variation in the incidence of Staphylococcus aureus bacteraemia in Auckland

1.1 Introduction

*Staphylococcus aureus* [SA] has the ability to cause a range of devastating diseases, in people of all ages and from all walks of life. SA is the primary bacterial threat to New Zealanders. In the late 1990s, the incidence of SA bacteraemia [SAB] in adults in New Zealand was approximately 41/100,000 \(^1\). In children, under the age of 16 years, the incidence was estimated to be 16.9/100,000 \(^2\). In comparison, the incidence of invasive pneumococcal disease in New Zealand in 2008 was 14.8/100,000, but the rates were over 60/100,000 in children under two years of age \(^3\). The peak incidence of meningococcal disease during New Zealand’s recent epidemic was 17.4/100,000 in 2001 \(^4\).

Previous reports have suggested that invasive disease caused by SA is more common in some ethnic groups living in New Zealand \(^1,2,5,6\). In a prospective study of people, Hill et al found that the relative risk of SAB for Māori was almost twice that of NZ European people; for Pacific people the relative risk of SAB was four times higher than NZ European people \(^1\).

The reasons for the variation in the incidence of disease caused by SA in different ethnic groups are not known, but deprivation is likely to be a contributing factor. Social deprivation can lead to poor health in a number of ways. Environmental exposure to cigarette smoke and household crowding were linked to an increased incidence of meningococcal disease in New Zealand children \(^7\). Social deprivation can also lead to inequalities in access to healthcare and in the quality of the healthcare that is received \(^8\).

Huggan et al, explored demographic variables related to the incidence of SAB in New Zealand \(^9\). They found that deprivation was associated with an increased risk of SAB; but they did not find an
increased incidence of SAB in Māori or Pacific people. However, this study might not have adequate power to detect a difference between ethnic groups due to the low numbers of Māori and Pacific people who live in Christchurch. It is important to note that their study was performed retrospectively and deprivation could only be estimated using population data.

A search of the literature did not find any reports of prospective studies of the effects of ethnicity or deprivation on the incidence of invasive disease caused by SA. In our setting this is of high importance; improved understanding of factors contributing to the differences in incidence is fundamental to future attempts at disease prevention.

This Chapter provides the results of the first large prospective study of adults with SAB performed to determine the impact that ethnicity and deprivation have on the incidence of SAB and its characteristics.
1.2 Materials and Methods

The Northern Y ethics committee of the NZ Ministry of Health provided ethical approval for this study.

1.2.1 Selection of cases of SAB

SAB was selected as the invasive disease model for this thesis. SAB was common in Auckland, and encompassed a wide range of disease entities in a wide range of human hosts. Furthermore, the diagnosis of SAB is usually straightforward, more than 90% of episodes can be detected by a single blood culture; and most episodes represent true bacteraemia.

A case was defined as a consenting adult [≥ 15 years] admitted to Middlemore Hospital or Auckland City Hospital after 26<sup>th</sup> February 2007, who had clinically significant bacteraemia caused by SA grown from a blood culture obtained by venipuncture or from a central venous catheter. Polymicrobial infections were not included. The clinical significance of each episode of SAB was based on clinical parameters, results of other investigations, and response to treatment. All cases were prospectively reviewed within 3 days of the onset of SAB, except for cases who died prior to review. The clinical records of cases who died were reviewed and, if required, the attending medical team were interviewed, after consent was obtained from the next of kin.

The laboratory staff contacted the investigator directly when a blood culture became positive; the laboratory records were also reviewed twice a week, at both study sites, to ensure that all cases had been captured. Review of the laboratory records for the years preceding the current study indicated that approximately 200 cases [including children] of SAB were identified at the two study sites each year. All consecutive adults with SAB were reviewed until 150 adult participants had been recruited.

1.2.2 Clinical review

The investigator, an infectious disease physician, performed all study procedures. Cases were contacted within the hospital and the study was explained using an interpreter, if required. Informed consent was obtained from the patients themselves, or from their next of kin, if the patient was unable to do so. On the single occasion that a patient was not admitted to hospital, contact was made by telephone, and the case was reviewed in his home. Clinical and demographic details were gathered during an interview with the case or their next of kin.

A short medical history was supplemented by information obtained from the patient’s hospital record. Information about the current infection, recent illnesses, past staphylococcal infection and past hospital admissions was obtained. All participants were asked specifically about a recent history of skin and soft tissue infection.
Demographic information including personal income, education level, and ethnicity was obtained using the questions employed in the 2006 NZ Census. The occupation and main source of income for each case was selected from one of ten categories by mutual agreement between the case and the investigator. The ten categories were: business, healthcare, homecare, professional, retired, sickness, student, tradesman, unemployed and unskilled.

Finally, data were collected about the household. This included an estimate of household income and basic demographic features of household members, as well as information about the home itself.

1.2.3 Methods of analysis and definitions used

Ethnicity

Ethnicity was self-reported, and participants were asked to select their ethnic group[s] from the same list used in the 2006 Census. In situations when participants selected more than one ethnic group, their ethnicity was reclassified as Māori, Pacific, or other, using prioritisation defined by the Ministry of Health. The study population consisted of 150 cases of SAB. For a number of comparisons cases were placed into three ethnic groups: Māori, Pacific people and people of all other ethnicities.

Occupation and main source of income

The sickness category included cases whose main source of income was a sickness or invalids benefit. Unemployed cases were those who received an unemployment benefit and were actively seeking work; unskilled cases were those who were employed in jobs that required no formal training. The homecare category included cases who chose not to work or whose unpaid occupation was the care of children or relatives.

Comparison with the Auckland population

In order to obtain a sample of SAB cases representative of all cases within the total population of Auckland, the study population came from two district health board catchments: Counties Manukau and Auckland. Comparison was made with population data from the 2006 NZ census for the two DHB catchments.

The NZ index of deprivation

The NZ index of deprivation [NZDep2006] was used to estimate the relative deprivation of the neighbourhood of residence of each case. The home address of each case of SAB was used to determine the residential block that they lived in; there are over 40,000 residential blocks for the population of NZ. The NZDep2006 score for each residential block was obtained from the
published values. The NZDep2006 index is an ordinal score of 1 to 10 and was derived for small residential blocks using data obtained from the 2006 Census \(^{14}\). A score of 10 indicates the most deprived areas. Three examples of the ten parameters used to derive the index for each block are people receiving a means tested benefit, people living in households with household income below a threshold, and people not living in their own home.

*Alcohol intake and cigarette smoking*

A history of excessive alcohol intake was defined as intake of more than 21 standard drinks per week in males and more than 14 standard drinks per week in females. These limits were the safe levels of intake advised by the Alcohol Advisory Council of New Zealand [http://www.alac.org.nz/](http://www.alac.org.nz/). A case was defined as a current cigarette smoker if they smoked more than one cigarette per day.

*Household crowding*

Household crowding was defined in two ways.

A bedroom occupancy rate was calculated from the ratio of the number of household occupants resident on at least three nights per week to the number of bedrooms. This definition was used for comparison with published values for Auckland City and Manukau City \(^{15}\).

The shared bedroom rate was the proportion of households where occupants, who were not couples, or children under the age of 15, shared one or more bedrooms.

*Access to healthcare*

Cases were asked if they were registered with a regular general practitioner’s [GP] surgery and whether their GP had been visited within the month or three months prior to the onset of SAB. They were also asked if they had a Community Services card. In New Zealand, a Community Services card is available on a means tested basis and provides the holder with benefits such as a reduction in healthcare costs.

*Past History of SA infection*

A past history of infection caused by SA was defined as any report of disease due to SA documented in the clinical records at any time in the past or a history of classical staphylococcal skin or soft tissue infection [e.g. skin boils] within the year prior to the episode of SAB.

*Acquisition of SA*

Place of acquisition was determined using accepted definitions \(^{16, 17}\). Community-onset disease was any illness when symptoms of the infection began in the community and was not associated with hospital admission in the past 12 months.
Community-onset healthcare-associated infection included cases who had recent healthcare exposure, including dialysis or residence in a long-term care facility or the presence of a permanent percutaneous device or indwelling catheter. For example, SA infection of an arterio-venous fistula leading to SAB in a person who received their haemodialysis at home was defined as community onset-healthcare associated infection.

Hospital-onset infection was an infection where the symptoms were not apparent until more than 48 hours after hospital admission, and the admission was for some other reason.

**Disease severity**

Disease severity was primarily measured using the acute physiology and chronic health evaluation score [APACHE II] \(^{18,19}\). The APACHE II score is a composite score used to determine the severity and prognosis of serious illness. A final score is formulated from clinical parameters, investigations and historical factors at the time of illness. The APACHE II score has been correlated with poor outcome in sepsis caused by SA \(^{13}\). In this thesis, the APACHEII score was calculated using parameters obtained on the day that the first positive blood culture was collected.

The maximal value of C-reactive protein [CRP] was also recorded. CRP is an acute phase protein that has not been validated as a prognostic indicator in SAB, but has prognostic value in left sided endocarditis \(^ {20}\). In this thesis, the highest value of CRP that was obtained in the first seven days after the day that the first positive blood culture was collected was recorded.

**Outcome of SAB**

Cure was survival without relapse for 12 months following the onset of SAB.

Relapse was any culture confirmed illness caused by SA in the six months following the cessation of treatment for the initial episode of SAB.

Death due to SA was a death that occurred as a direct result of SAB. Death not directly due to SA was defined as a death that occurred in a case of SAB after the signs and symptoms of sepsis had resolved, and when an alternative cause for death was more likely.

**1.2.4 Statistical methods**

Analysis was performed using Microsoft Excel 2008 for Macintosh [Seattle, USA]; and SPSS 16.0 [Chicago, USA].

The Kolmogorov-Smirnov test was used to test differences between the empirical distribution functions of the ages of cases in each ethnic group. The one sample t-test was used to compare the sample mean values of categories with the expected mean obtained from population census
data. The $\chi^2$ test and Fisher’s exact test were used to determine the statistical significance of comparisons of categorical data. $\chi^2$ values and the degrees of freedom [d.f.] were recorded with the $P$ value. The ANOVA test was used to compare means of continuous data; Fisher’s value [F] and the d.f. were recorded with each $P$ value. The Mann-Whitney U test was used to compare variables that were not normally distributed. The Mantel-Haenszel test of conditional independence was used to examine the effect of age on differences between ethnic groups.
1.3 Results

150 cases of SAB were recruited to the study over 11 months. Six other people met the case definition, but were not recruited into the study. Two declined to take part; two died prior to contact being made with them or their next of kin; and two were discharged from hospital prior to contact with the investigator, and could not be contacted subsequently.

1.3.1 Age, gender and ethnicity of cases of SAB

Cases with SAB [median age 56, interquartile range 40-73] were older than adults in the Auckland population [median age 34 years]. Only one third [48/150, 32%] of the cases were female. Females [median age 58, IQ range 38-79] and males [median age 56, IQ range 44-69] did not differ significantly in age [Mann-Whitney U test, \( P=0.426 \)].

The annualised incidence of SAB was 20/100,000 population members. 52/150 [35%] of the cases were Pakeha, which gave an incidence of SAB of 14/100,000 Pakeha population members. 11/150 [5%] of the cases were Asian [includes 3/150, 2% Indian cases]; 9/150 [7%] were of other European ethnicity; and 2/150 [1%] were Middle Eastern. These 74 cases were grouped into cases of other ethnicities.

The SAB annualised incidence for Māori [36/100,000] and Pacific people [34/100,000] was higher than the incidence for people of other ethnicities. Almost half [21/44, 47%] of the Pacific cases with SAB were Samoan; and the incidence of SAB was similar among the three largest populations of Pacific people living in Auckland: Samoan, Tongan and Cook Island Māori [Table 1.1]. The variance in incidence in other ethnic groups of Pacific people should be interpreted with caution because of the small numbers of cases in each group. The proportion of females did not differ significantly between each ethnic group [Māori cases 10/32, 31%; Pacific cases 15/44, 34%; cases of other ethnicities 23/74, 31%].
Māori and Pacific cases of SAB were younger than cases of other ethnicities. The median age of Māori and Pacific cases was not significantly different [Māori 48 years, IQ range 35-60; Pacific people 46, IQ range 34-57], however, the median age of the cases of other ethnicities was almost 25 years older [70 years, IQ range 51-80]. Figure 1.1A shows the cumulative distribution function of age for each ethnic group. The number of cases rose steadily with age for the Māori and Pacific cases. The number of cases under the age of 65 years was similar between the three ethnic groups, but the number of cases of other ethnicities rose rapidly in old age [Figure 1.1B].

The incidence of SAB increased with age in all three ethnic groups, but reached a maximum rate in Pacific people between the ages of 55 to 64, in Māori and people of other ethnicities the maximum rate was reached between the ages of 75 and 84. There was a significant difference between the age distributions of Māori or Pacific cases and cases of other ethnicities [Kolmogorov-Smirnov test; cases of other ethnicities versus Māori $P=0.006$; cases of other ethnicities versus Pacific cases $P=0.011$]; there was no difference in the age distribution between Māori and Pacific cases. In all age groups under the age of 65 years the incidence of SAB was higher in Māori and Pacific people than people of other ethnicities [Figure 1.1C]. The incidence of SAB in people of other ethnicities increased after the age of 65. The incidence if SAB also increased sharply in Māori over the age of 75, but there were few cases of SAB at this age group, and the high incidence is a reflection of the small number of elderly Māori in the total population of Auckland.

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Number</th>
<th>Incidence/100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>Pacific people</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>Samoan</td>
<td>21</td>
<td>34</td>
</tr>
<tr>
<td>Tongan</td>
<td>9</td>
<td>29</td>
</tr>
<tr>
<td>Cook Island</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Fijian</td>
<td>4</td>
<td>113</td>
</tr>
<tr>
<td>Pacific Other</td>
<td>4</td>
<td>86</td>
</tr>
<tr>
<td>Niuean</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>74</td>
<td>14</td>
</tr>
<tr>
<td>Pakeha</td>
<td>52</td>
<td>14</td>
</tr>
<tr>
<td>Asian</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>Other European</td>
<td>9</td>
<td>NA</td>
</tr>
</tbody>
</table>

1 The number of people of Other European ethnicities in the total population of Auckland was not available [NA]

Table 1.1. The ethnicity and annualised incidence of 150 cases of SA bacteraemia by ethnic group.
Figure 1.1 [A] The cumulative distribution function of age for Māori, Pacific and Other cases of SAB; [B] the number of cases of SAB by ethnic group; [C] the incidence/100,000 of SAB by age and ethnic group.
1.3.2 The education level, personal income and occupation of cases of SAB

6/150 [4%] of the cases, all aged over 65 years, had not attended secondary school. The proportion of Pacific cases with a tertiary qualification [8/44, 18%] was not significantly different from the proportion of cases of other ethnicities [18/74, 24%], and was not significantly different from total Auckland population in 2006 [18%]. Only one Māori case had a tertiary qualification [1/32, 3%] and this was lower than the proportion in the total population of Auckland [18%, one-sample T-test, \( P < 0.001 \)].

Four cases [4/150, 3%] were unemployed [and looking for work] and this was not significantly different from the rate in the total population of Auckland [4% unemployed]. The main source of income for the cases in each ethnic group is shown in Figure 1.2. Over half of the cases of other ethnicities [38/74, 51%] were retired, which was higher than the proportion of retired Māori cases [3/32, 9%] and Pacific cases [6/44, 14%] \( \chi^2 = 27.8, 2 \text{ d.f.}, P < 0.08 \).

Figure 1.2. The main source of income [%] for cases of SAB by ethnic group.
The proportion of the cases who received the sickness or invalids benefit [22/150, 15%] was higher than the proportion in the total population of Auckland [4%, one sample t test, \( P<0.01 \)]. The proportion of Māori cases [5/29, 17%] who received the sickness benefit, excluding those retired and over the age of 65 years [who were not eligible for the benefit], was not significantly different from the proportion Pacific cases [9/38, 24%] or cases of other ethnicities [8/36, 22%] who received the sickness benefit.

89/150 [59.3%] of the cases earned less than $20,000 in the year preceding their episode of SAB; this proportion was greater than the total population of Auckland [36%, one sample t test, \( P<0.01 \)]. The proportion of cases who earned less than $20,000 did not differ significantly between Māori [22/32, 69%, 51-82], Pacific [32/44, 73%, 58-84] and other ethnicities [35/74, 47%, 36-59] \( [\chi^2 3.69, 2 \text{ d.f., } P=0.16] \). The proportion of cases who earned more than $50,000 [15/150, 10%, 6-16] was less than the total population of Auckland [19%, one sample t-test, \( P<0.01 \)].

The maximum amount that a person receiving a sickness or invalids benefit in 2007 could earn was less than $20,000 [before tax, for a sole parent on the invalids benefit] and the maximum amount that an individual could earn from publicly funded NZ superannuation was also less than $20,000 [before tax]. The proportion of cases who had not retired and who did not receive a sickness or invalids benefit that earned less than $20,000 [76/150, 51%, 43-59] did not differ significantly from the proportion of the total population of Auckland who earned less than $20,000 [36%].

### 1.3.3 The deprivation index of the neighbourhood of domicile of SAB cases

The NZDep2006 score is a composite score used to estimate the deprivation of a household based on census data for the neighbourhood. Half of the Māori [15/30, 50%] and Pacific [23/44, 52%] cases of SAB lived in the most socially deprived areas of Auckland; the NZDep2006 score of their residential address was 10. In contrast, only one in ten cases of other ethnicities [7/71, 10%] lived in areas of Auckland with NZDep2006 scores of 10. In the total New Zealand population 24% of Māori, 36% of Pacific people and 6% of people of other ethnicity live in areas with NZDep2006 scores of 10.²¹

When the deprivation scores were grouped according to areas of the least deprivation [score 1, 2 or 3]; areas of modest deprivation [score 4, 5 or 6]; and areas of highest deprivation [score 7, 8, 9 and 10] the differences between ethnic groups became more marked [Figure 1.3]. 40/44 [91%] Pacific cases lived in areas of highest deprivation, compared with 25/74 [39%] cases of other ethnicities \( [\chi^2 16.36, 2 \text{ d.f., } P<0.001] \).
1.3.4 Alcohol usage and cigarette smoking

15/32 [47\%] of Māori cases reported excessive alcohol consumption as did 34/74 [46\%] of cases of other ethnicities, which was not significantly different to the proportion of Pacific cases [11/44, 25\%] who reported excessive alcohol consumption.

38/150 [25\%] of the cases smoked regularly, there was a non significant trend towards higher rates of smoking in Māori [12/32, 38\%] and Pacific cases [14/44, 32\%] than in cases of other ethnicities [12/74, 16\%] [χ² 5.4, 2 d.f., P=0.06].

The number of cases who reported regular use of other recreational drugs was too small to allow comparisons to be drawn.

1.3.5 The household characteristics of cases of SAB

The proportion of cases who owned their place of residence [71/150, 47\%] was not significantly different from that in the total population of Auckland [55\%, one sample t-test P=0.063]. However, Māori [7/32, 22\%] and Pacific cases [10/44, 23\%] were less likely to own their home than cases of other ethnicities [54/74, 73\%] [χ²=20.3, 2 d.f., P<0.01]. The number of cases who were renting
their home from Housing New Zealand was higher in Māori [11/32, 34%] and Pacific cases [21/44, 48%] than in cases of other ethnicities [4/74, 5%] \( \chi^2 = 30, 2 \text{ d.f.}, P < 0.01 \).

The household size of Pacific cases [average number of members 5.0 s.d. 1.8] was larger than the household size of either Māori [3.8 s.d. 2.0] or cases of other ethnicities [2.4 s.d. 1.5] [ANOVA, F=33, 2 d.f., P<0.01]. Māori [1.23] and Pacific cases [1.69] had higher bedroom occupancy rates than cases of other ethnicities [0.88]. The corresponding value for Auckland City was 0.96 and for Manukau City was 1.03. Sharing of a bedroom was more common in the households of Māori [6/31, 19%] and Pacific cases [22/44, 50%] than in the households of cases of other ethnicities [5/70, 7%] \( \chi^2 = 28, 2 \text{ d.f.}, P < 0.01 \).

A high proportion of households [37/147, 25%] did not use any form of home heating during winter. The households of Māori [9/31, 29%] and Pacific cases [22/44, 50%] were less likely to use any form of home heating than the households of cases of other ethnicities [6/72, 8%] \( \chi^2 = 19.1, 2 \text{ d.f.}, P < 0.01 \).

### 1.3.6 Access to healthcare

A high proportion of cases [138/150, 92%] were registered with a GP, and this was not significantly different between ethnic groups; although cases of other ethnicities were more likely to have seen a GP in the three months prior to their episode of SAB [Table 1.2]. The proportion of cases, who held a community services card, was not significantly different between the three ethnic groups.

<table>
<thead>
<tr>
<th>Ethnicity [n]</th>
<th>GP</th>
<th>GP within 1 month</th>
<th>GP within 3 months</th>
<th>Community services Card</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori [32]</td>
<td>29 [91%]</td>
<td>15 [47%]</td>
<td>18 [56%]</td>
<td>24 [75%]</td>
</tr>
<tr>
<td>Pacific [44]</td>
<td>39 [89%]</td>
<td>21 [48%]</td>
<td>28 [64%]</td>
<td>27 [61%]</td>
</tr>
<tr>
<td>Other [74]</td>
<td>70 [95%]</td>
<td>40 [54%]</td>
<td>64 [87%]</td>
<td>48 [65%]</td>
</tr>
</tbody>
</table>

\( \chi^2 = 8.4, 2 \text{ d.f.}, P = 0.015 \)

Table 1.2. Access to a regular GP and visits in the month and three months prior to SAB.

### 1.3.7 Past history of infection caused by SA and comorbid illness

Comorbid illnesses, such as diabetes, are well known risk factors for disease caused by SA and are likely to confound ethnic variation in the incidence of SAB. Diabetes was more common in Pacific cases [24/44, 54%] than Māori cases [12/32, 38%] or cases of other ethnicities [16/74, 22%] \( \chi^2 = 9.3, 2 \text{ d.f.}, P < 0.01 \).
The mean body mass index \([\text{BMI}, \text{weight [kg]} \div \text{height squared [m}^2\text{]}]\) of Māori [27.3 s.d. 5.4] was not significantly different from the mean BMI of cases of other ethnicities [25.0 s.d. 5.7] and in both groups there were four cases with a BMI of greater than 35. Pacific cases had a higher mean BMI [32.0 s.d. 8.4, ANOVA \(F= 15.74, 2 \text{ d.f.}, P<0.01\)] and there were 16 cases with a BMI greater than 35. When the cases with the highest BMI [\(\geq 35\)] were removed from analysis, the mean BMI was not significantly different between ethnic groups.

There was a difference in the proportion of cases who had end stage renal failure [ESRF] requiring renal replacement therapy between ethnic groups. The proportion of Pacific cases who had ESRF [20/44, 45%] was greater than the proportion of Māori cases [11/32, 24%] and cases of other ethnicities [12/74, 16%] who had ESRF \([\chi^2=12.184, 2 \text{ d.f.}, P<0.01]\.\)

Haemodialysis performed via a fistula or central venous line [CVL] was an important contributing factor in Pacific and Māori cases. 7/32 [22%] of Māori cases and 9/44 [21%] of Pacific cases were receiving haemodialysis via a CVL compared with 3/74 [4%] of cases of other ethnicities \([\chi^2=7.75, 2 \text{ d.f.}, P=0.02]\.\)

Ischaemic heart disease [IHD] was common among cases of SAB; 55/150 [37%] had a history of myocardial infarction or coronary artery angioplasty or bypass graft surgery. The proportion of cases who had IHD was similar in all ethnic groups [Māori 12/32, 38%; cases of other ethnicities 31/74, 42%; Pacific cases 12/44, 27%; \(\chi^2=2.55, 2 \text{ d.f.}, P=0.28\)]. Other common illnesses, such as cancer and asthma were less prevalent among cases with SAB, and were not significantly different between ethnic groups.

1.3.8 Past History of infection caused by SA

38/150 [25%] of cases of SAB had a past history of SA infection, but there were no significant differences between ethnic groups. Only 12/150 [8%] reported a history of SA infection that immediately preceded their episode of SAB. All of these cases had SAB secondary to skin and soft tissue infection.

1.3.9 Acquisition of SA infection

61/150 [41%] of the cases had community-onset infection. This proportion did not differ significantly between ethnic groups \([\chi^2=4.56, 2 \text{ d.f.}, P=0.102] [\text{Table 1.3}\.\)] 9/32 [28%] Māori cases and 15/44 [34%] Pacific cases had community-onset healthcare-associated SAB, which was greater than the proportion of cases of other ethnicities that had community-onset healthcare-associated SAB [8/74, 11%, \(\chi^2=10.03, 2 \text{ d.f.}, P<0.01]\. Most of the Pacific community-onset healthcare-associated cases of SAB [12/15, 80%] represented infections associated with arteriovenous fistulae or central venous lines required for haemodialysis. A smaller proportion of Māori [5/9, 56%] and other [4/8, 50%] cases had community-onset healthcare-associated SAB
associated with haemodialysis. The proportion of Māori and Pacific cases who suffered hospital onset SAB was less than cases of other ethnicities [$\chi^2 = 13.8, 2$ d.f., $P<0.01$]; however, the incidence/100,000 of hospital onset infection was not different between ethnic groups.

<table>
<thead>
<tr>
<th>Ethnicity [n]</th>
<th>community-onset</th>
<th>community-onset incidence</th>
<th>community-onset healthcare associated</th>
<th>hospital-onset</th>
<th>hospital-onset incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori [32]</td>
<td>17 [53%]</td>
<td>19</td>
<td>9 [28%]</td>
<td>10</td>
<td>6 [20%]</td>
</tr>
<tr>
<td>Pacific [44]</td>
<td>20 [46%]</td>
<td>15</td>
<td>15 [34%]</td>
<td>11</td>
<td>9 [21%]</td>
</tr>
<tr>
<td>Other [74]</td>
<td>24 [32%]</td>
<td>4</td>
<td>8 [11%]</td>
<td>1</td>
<td>42 [57%]$^1$</td>
</tr>
<tr>
<td>Total [150]</td>
<td>61 [41%]</td>
<td>8</td>
<td>32 [21%]</td>
<td>4</td>
<td>57 [38%]</td>
</tr>
</tbody>
</table>

$\chi^2 = 13.8, 2$ d.f., $P<0.01$

Table 1.3. The ethnic group of cases of SAB and annualised incidence/100,000 by place of onset.

### 1.3.10 The cause of SAB: diagnosis and MRSA

SAB was associated with a diverse range of infections [Figure 1.4]. The most common cause of SAB was infection related to invasive devices. These caused over one-third of the infections [57/150, 38%] and over half of these were caused by infection of peripheral venous cannulae [29/57, 51%]. Infections of central venous catheters caused 21/57 [37%] of SAB episodes and 14/21 [67%] of these represented infections of haemodialysis catheters. Infection of permanent pacemakers was the most common [5/7, 71%] of the other causes of infections involving invasive devices.

Skin and soft tissue infections were the cause of SAB in 33/150 [22%]. 19/33 [58%] skin and soft tissue infections began in the community; 5/6 healthcare associated infections occurred in cases who received haemodialysis, and three of these cases had infection related to their arterio-venous fistulae. All 8/8 of the hospital onset cases of SAB secondary to skin and soft tissue infection had infection of surgical wounds, four of these occurred following caesarean section.

The number of episodes of SAB caused by MRSA was low [11/150, 7%] and was not different between ethnic groups. Māori [1/32, 3%] and Pacific cases [6/44, 14%] had lower rates of SAB caused by peripheral venous cannulae than cases of other ethnicities [22/74, 30%; $\chi^2 = 9.243, 2$ d.f., $P<0.01$]. There were no significant differences between ethnic groups for any of the other diagnostic categories.
1.3.11 Disease severity and length of hospital stay

The median APACHEII score of cases on the day that the first positive blood culture was obtained was 12, IQ range 7-16. This was not significantly different between cases from the three different ethnic groups [Mann Whitney U test, $P=0.426$]. Likewise the highest CRP value [mean 135, s.d. =110] obtained in the first seven days after the onset of SAB was not significantly different between ethnic groups [Mann Whitney U test, $P=0.868$].

The average length of hospital stay of the 133 cases, who did not die during their inpatient stay, was 19 s.d.17 days. The mean length of stay for Māori cases was 14 days, s.d. =8; the mean length of stay for Pacific cases was 18 days, s.d. =19; and the mean length of stay for cases of other ethnicities was 23 days, s.d. =17. The difference between Māori and Pacific cases was not significant when the six cases who stayed longer than 2 months were removed from analysis, but both ethnic groups had shorter durations of hospital stay than cases of other ethnicities [ANOVA $F=6.515, P=0.002$]. The length of stay was not different between cases treated at Auckland Hospital or Middlemore Hospital.

1.3.12 The outcome and mortality of SAB

118/150 [79%] cases were cured; there were no differences in the cure rate between ethnic groups. 17/32 [53%] who were not cured, died during their inpatient stay. A further 10 cases [10/32, 31%] suffered a relapse and 5/32 [16%] died during the 12 months of follow-up.
22/150 [15%] cases died as a direct result of their SAB. The median age of these cases was 61 years, IQ range 55-79, which was similar between ethnic groups [Māori 61 years, IQ range 61-79; Pacific 56 years, IQ range 47-61; other ethnicities 71 years, IQ range 58-83]. The distribution of ages of the cases who died within each ethnic group did not differ significantly [Kolmogorov-Smirnoff test, \( P=0.107 \)]. The proportion of cases who died did not differ significantly between the different ethnic groups [Māori 5/32, 16%; Pacific 5/44, 11%; other ethnicities 12/74, 16%].

There were no significant differences between those who died and those who survived with regard to the mean NZDep2006 scores; the proportion of cases registered with a GP; the proportion of cases with a community services card; the proportion of cases who received a sickness benefit; or the proportion of cases who earned less than $20,000. 10/57 [17%] of cases with nosocomial infection died, which was not significantly different from the proportion of cases with community–onset infection [9/61, 15%] or healthcare-associated SAB [3/32, 9%].

There were no significant differences in comorbidities between those who died and those who survived. For example 9/22 [41%] of those who died were diabetic, which was not significantly different from the total number of cases [52/150, 35%] [Fisher’s exact test, \( P=0.6 \)]. Likewise there were no significant differences in rates of cigarette smoking and excessive alcohol intake.

The APACHEII score on the day of onset of SAB did correlate with mortality. The median score was 16, IQ range 13-19, in those who died, compared to 9, IQ range 6-14, in those who survived [Mann-Whitney U test, \( P<0.001 \)]. The peak value of CRP within 7 days of onset of SAB was not different between those who died or survived.

The time from onset of bacteraemia to death is shown in Figure 1.5. Despite the low number of cases who died, there were two modes evident, those who died early [<14 days] and those who died late [>28 days]. The majority [15/22, 68%] died within 14 days of the onset of bacteraemia, however, all but one of the other seven deaths occurred more than 28 days after the onset of SAB.
1.3.13 Comparison of indicators of deprivation with population values

Comparable population values of six potential indicators of deprivation were available for comparison with the values obtained from cases of SAB. Compared with the total population of Auckland, cases of SAB were more likely to earn less than $20,000 per year, receive a sickness or invalids benefit, and smoke cigarettes [Table 1.4]. Cases of SAB were less likely to earn over $50,000 per year, but there was no significant difference in the proportion of cases who had a tertiary qualification or who owned their home compared with the total population of Auckland.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>number [%], 95%CI</th>
<th>population proportion</th>
<th>P value$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>income &lt; $20,000/year</td>
<td>89 [59, 32-47]</td>
<td>36%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>income &gt; $50,000/year</td>
<td>15 [10, 6-16]</td>
<td>19%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>sickness/invalids benefit</td>
<td>22 [15, 10-21]</td>
<td>4%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>cigarette smoking</td>
<td>38 [25, 19-33]</td>
<td>17%</td>
<td>0.021</td>
</tr>
<tr>
<td>home owner</td>
<td>71 [47, 40-55]</td>
<td>55%</td>
<td>0.063</td>
</tr>
<tr>
<td>tertiary qualification</td>
<td>27 [18, 13-25]</td>
<td>18%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

$^1$ one sample t test

Table 1.4. Indicators of deprivation amongst 150 cases of SAB compared with population proportions.

1.3.14 Summary of ethnic variation among cases of SAB

Table 1.5 shows a summary of the characteristics of cases of SAB of Māori, Pacific and other ethnicities. The majority of statistically significant differences were in indicators of deprivation, which differed between cases by ethnicity. Diabetes and end-stage renal failure were more common in Pacific cases of SAB. Consequently, cases of other ethnicities were less likely to receive haemodialysis via a central venous catheter.
<table>
<thead>
<tr>
<th></th>
<th>Māori</th>
<th>Pacific</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>32</td>
<td>44</td>
<td>74</td>
</tr>
<tr>
<td>Incidence of SAB</td>
<td>33/100,000</td>
<td>30/100,000</td>
<td>12/100,000</td>
</tr>
<tr>
<td>Mean age, years [s.d.]</td>
<td>48 [18]</td>
<td>46 [16]</td>
<td>65 [20]*</td>
</tr>
<tr>
<td>Female</td>
<td>10 [31, 18-49]</td>
<td>15 [34, 22-49]</td>
<td>23 [31, 22-42]</td>
</tr>
<tr>
<td>Tertiary qualification</td>
<td>1 [3, 0-17]*</td>
<td>8 [18, 9-32]</td>
<td>18 [24, 16-35]</td>
</tr>
<tr>
<td>Unemployed</td>
<td>0 [0, 0-13]</td>
<td>4 [9, 3-22]</td>
<td>0 [0, 0-6]</td>
</tr>
<tr>
<td>Retired</td>
<td>3 [9, 3-25]</td>
<td>6 [14, 6-27]</td>
<td>38 [51, 40-62]*</td>
</tr>
<tr>
<td>Income &lt;$20,000/annum</td>
<td>22 [69, 51-82]</td>
<td>32 [73, 58-84]</td>
<td>35 [47, 36-59]</td>
</tr>
<tr>
<td>Income &lt;$20,000/annum*</td>
<td>14 [41, 28-61]</td>
<td>16 [36, 24-51]</td>
<td>9 [12, 6-22]*</td>
</tr>
<tr>
<td>Excess alcohol intake</td>
<td>15 [47, 31-64]</td>
<td>11 [25, 14-40]</td>
<td>34 [46, 35-57]</td>
</tr>
<tr>
<td>Mean household size [s.d.]</td>
<td>3.8 [2.0]</td>
<td>5.0 [1.8]</td>
<td>2.4 [1.5]</td>
</tr>
<tr>
<td>Bedroom occupancy rate</td>
<td>1.23</td>
<td>1.69</td>
<td>0.88</td>
</tr>
<tr>
<td>Shared bedroom</td>
<td>6 [19, 9-37]</td>
<td>22 [50, 36-64]</td>
<td>5 [7, 3-16]*</td>
</tr>
<tr>
<td>No home heating</td>
<td>9 [29, 16-47]</td>
<td>22 [50, 36-64]</td>
<td>6 [8, 4-17]*</td>
</tr>
<tr>
<td>Registered with GP</td>
<td>29 [91, 75-98]</td>
<td>39 [89, 76-96]</td>
<td>70 [95, 87-98]</td>
</tr>
<tr>
<td>GP visit &lt; 3 months</td>
<td>18 [56, 39-72]</td>
<td>28 [64, 49-76]</td>
<td>64 [87, 77-93]*</td>
</tr>
<tr>
<td>CSC</td>
<td>24 [75, 58-87]</td>
<td>27 [61, 47-74]</td>
<td>48 [65, 53-75]</td>
</tr>
<tr>
<td>Diabetic</td>
<td>12 [38, 23-55]</td>
<td>24 [54, 40-68]*</td>
<td>16 [22, 14-33]</td>
</tr>
<tr>
<td>SAB due to PVC infection</td>
<td>1 [3, 0-17]</td>
<td>6 [14, 6-22]</td>
<td>22 [30, 20-41]*</td>
</tr>
</tbody>
</table>

1. sickness or invalids benefit; not including cases who have retired or >65 years
2. not including cases who have retired or receive a sickness or invalids benefit
3. residential address not applicable in 2 Māori and 3 other cases of SAB
4. for those who stayed < 2 months
5. X² test, P<0.05
6. HNZ, housing New Zealand; GP, general practitioner; CSC, community services card; BMI, body mass index [kg m⁻²]; ESRF, end-stage renal failure; HD, haemodialysis; CVL, central venous line; IHD, ischaemic heart disease; PVC, peripheral venous cannula

Table 1.5 Characteristics [number [%, 95% CI] unless otherwise stated] of cases of SAB by ethnicity. The boxes shaded grey indicate significant differences between ethnic groups.
1.3.15 The difference between ethnic groups after stratification by age

The number of elderly Māori and Pacific cases of SAB was small [Figure 1.1B] and so cases were stratified into two groups according to the median age, 56 years. Only 12 Māori cases and 11 Pacific cases were aged more than 56 years, compared with 51 cases of other ethnicities. For this reason, and because of the similarity of the selected characteristics between Māori and Pacific cases [Table 1.5], they were grouped and compared with cases of other ethnicities. Following this stratification, several differences were no longer significant, including holding a tertiary qualification, hospital-onset SAB, retirement from their occupation and a visit to the GP within the last three months. Table 1.6 shows the Mantel-Haenszel common odds ratio for Māori and Pacific cases compared to cases of other ethnicities.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-stage renal failure</td>
<td>4.9</td>
<td>2.0-11.7</td>
</tr>
<tr>
<td>Haemodialysis</td>
<td>5.4</td>
<td>1.9-15.0</td>
</tr>
<tr>
<td>Diabetes</td>
<td>5.0</td>
<td>2.2-11.5</td>
</tr>
<tr>
<td>Healthcare associated SAB</td>
<td>3.8</td>
<td>1.5-9.9</td>
</tr>
<tr>
<td>SAB associated with PVC</td>
<td>0.3</td>
<td>0.1-0.8</td>
</tr>
<tr>
<td>Personal income &lt;$20,000/annum</td>
<td>3.2</td>
<td>1.5-6.7</td>
</tr>
<tr>
<td>NZDep2006 score 7-10</td>
<td>7.3</td>
<td>3.2-16.4</td>
</tr>
<tr>
<td>Home owner</td>
<td>0.14</td>
<td>0.06-0.31</td>
</tr>
<tr>
<td>Home rented from housing New Zealand</td>
<td>4.9</td>
<td>2.0-11.7</td>
</tr>
<tr>
<td>Shared bedroom in the household</td>
<td>7.7</td>
<td>2.6-23.5</td>
</tr>
<tr>
<td>Home heating</td>
<td>0.2</td>
<td>0.09-0.52</td>
</tr>
</tbody>
</table>

PVC, peripheral venous cannula

Table 1.6. The odds ratios of significant characteristics [P<0.05], following age stratification, of Māori and Pacific cases of SAB compared with cases of other ethnicities.

1.3.16 The incidence of SAB adjusted for the presence of renal disease and diabetes

The annualised incidence/100,000 of SAB in Pacific people who did not have renal failure and/or diabetes [10/100,000] was similar to the incidence of SAB in people of other ethnicities who did not have renal failure or diabetes [9/100,000] [Table 1.7]. The incidence of SAB in Māori cases who did not have diabetes and/or renal failure was 18/100,000.
Table 1.7. The annualised incidence/100,000 population members of cases of SAB amongst Māori, Pacific and other ethnicities adjusted for the presence of end-stage renal failure (ESRF), diabetes or both.

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>no ESRF</th>
<th>no diabetes</th>
<th>no ESRF or diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori</td>
<td>36</td>
<td>24</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>Pacific</td>
<td>34</td>
<td>18</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>other</td>
<td>14</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>14</td>
<td>13</td>
<td>11</td>
</tr>
</tbody>
</table>

The significantly increased risk of healthcare-associated SAB in Māori and Pacific cases [Table 1.5] was no longer evident following adjustment for the presence of renal failure [adjusted $\chi^2 = 1.6$, 1 d.f., $P=0.206$, Mantel-Haenszel test of independence, $P=0.083$, odds ratio 0.15-1.01].

1.3.17 Factors associated with community-onset SAB

Community-onset SAB was more common in Māori and Pacific cases combined [37/76, 49%, 38-60] compared with cases of other ethnicities [24/74, 32%, 23-44, $\chi^2 = 4.1$, $P=0.043$] [Table 1.5]. Community onset-healthcare associated SAB was also more common in Māori and Pacific cases [Māori and Pacific 24/76, 32%, 22-43; other 8/74, 11%, 5-20; $\chi^2 = 9.64$, $P<0.01$]. Hospital-onset SAB was less common in Māori and Pacific cases [15/76, 20%, 12-30] than cases of other ethnicities [42/74, 57%, 45-67 $\chi^2 = 21.8$, $P<0.01$].

Social deprivation was also likely to impact upon the incidence of community-onset SAB. Cases with community-onset SAB were more likely to be under the age of 55 years, not be registered with a GP, not own their home, live in an area with NZDep2006 scores between 7-10, and were more likely to live in a household which contained adults who were not couples but shared a bedroom [Table 1.8].

Table 1.8. The statistically significant [$P<0.05$] differences between cases of SAB with [$n=61$] and without [$n=89$] community-onset.
1.4 Discussion

This study determined the demographic, social and clinical details of 150 consecutive adult cases of SAB at two large public hospitals in Auckland. Due to the prospective study design, detailed information was able to be collected about clinical and social factors, including information about each case’s household.

Māori and Pacific people have a high incidence of SAB compared with people of other ethnicities. The results of the current study indicate that the incidence of SAB does not appear to have increased over the last decade. In the late 1990’s the incidence of SAB in adults in Auckland and Christchurch was 41/100,000/year; and this incidence was higher than the reported incidences from countries in Europe and North America. The incidence of SAB in Christchurch between 1998 and 2006 was 22/100,000/year and was lowest in Māori people.

Unlike the report from Christchurch, the current study identified striking differences in the incidence of SAB between ethnic groups; in Māori people the incidence of SAB was 36/100,000 compared with 14/100,000 for people of other ethnicities.

In the current study, the higher incidence of diabetes and renal failure amongst Māori and Pacific cases of SAB accounted for most of the increased incidence of community—onset healthcare-associated SAB observed in these groups. When cases with diabetes and/or renal failure were excluded the incidence of SAB among Pacific people reduced to the same level as people of other ethnicities; for Māori people the incidence halved, but remained elevated compared with people of other ethnicities. Thus, diabetes and renal failure are important risk factors for the development SAB among Māori and Pacific people in Auckland. Diabetes and end-stage renal failure are important risk factors for SAB; in Auckland, in 2006/7 the prevalence of diabetes in Māori and Pacific people was twice that of NZ European people. Likewise the prevalence of renal failure requiring dialysis in adults between the ages of 45 and 64 was approximately ten times higher in Māori and Pacific people compared with people of other ethnicities.

Infections related to vascular access for haemodialysis accounted for most of the difference between ethnic groups in the incidence of community-onset healthcare-associated SAB. Eighty per cent of Pacific cases of community-onset healthcare associated SAB were associated with haemodialysis and when the analysis was stratified according to the presence or absence of renal failure the incidence of community-onset healthcare-associated SAB dropped to the same level as that found among cases of other ethnicities. The risk of SAB is particularly high in people who receive haemodialysis via a central venous catheter but is reduced by the use of an arterio-venous fistula. A review of the New Zealand dialysis registry found that Māori and Pacific people were less likely to have dialysis performed via an arterio-venous fistula.

Despite the finding that renal disease and diabetes could account for most of the ethnic variation in the incidence of SAB, important differences were found that could not be explained by confounding due to diabetes and/or renal failure. While the proportion of cases of community-
onset SAB in Auckland was within the range of those reported in other countries, the incidence of community-onset SAB was higher in Māori and Pacific cases. The proportion of community-onset cases of SAB was lower in the current study [41%] than the proportion found in New Zealand in the late 1990's [50%] 1. However, in that study, bacteraemia that developed in rest-home residents was considered to have been community-acquired. A study of SAB in Sweden that used the same definitions as the current study found that half of their cases of SAB were community-acquired 27; in contrast, in Minnesota between 1998 and 2005 only 18% of cases of SAB were community-acquired 22.

The impact of deprivation was most likely to be evident in community-onset cases, and further analysis was performed to compare the characteristics of community-onset cases of SAB with those who did not have community-onset SAB. Studies of staphylococcal disease epidemiology have seldom included measures of deprivation as potential risk factors for disease. Social deprivation was associated with an increased risk of SAB in Christchurch, but deprivation was solely estimated using NZDep2006 scores 9. In the United Kingdom, patients from areas of high deprivation had higher rates of surgical site infection caused by methicillin-resistant SA following cardiac surgery 26. In the current study, community-onset SAB cases were more likely to live in areas of the greatest deprivation and live in a household which includes adults who shared a bedroom but were not couples; community-onset SAB cases were less likely to be over the age of 55 years, own their home, be registered with a general practitioner and have co-morbid renal failure or cancer.

It is likely that deprivation also plays an important part in the development and progression of renal failure and diabetes, yet these illnesses were not associated with an increased risk of community-onset SAB. Access to healthcare was similar in diabetics in Auckland of different ethnic groups and between diabetics who lived in areas with higher NZDep2006 scores 21. Access to healthcare could be an important factor in the development of SAB; however, opportunities to prevent SAB amongst cases in the current study may have been limited to the few cases whose SAB was caused by skin and soft tissue infection and reported a preceding lesion consistent with early SA infection. The increased incidence of SAB observed in Māori and Pacific cases was not caused by reduced access to healthcare. The proportion of cases who were registered with a general practitioner and who had seen their general practitioner in the month prior to their episode of SAB was not different between ethnic groups.

In the current study many of the differences between ethnic groups were related to indicators of social deprivation. However, these differences were also a reflection of the societal differences between ethnic groups. In order to further examine whether social deprivation was a potential risk factor for the development of SAB, the indicators of social deprivation for cases of other ethnicities were compared with the total population of Auckland. These cases did not differ significantly from the total population of Auckland: 16% were cigarette smokers, 73% owned their home, and 7% of households contained two adults that shared a bedroom but were not couples.
When retired cases were taken into account the proportion of cases who earned less than $20,000 per year was not different to the total population.

A strength of the current study was the inclusion of a large number of cases of SAB and every effort was made to include all consecutive cases, including those who had died early in their hospital stay. The lack of a control cohort imposed limitations on the conclusions that could be drawn from comparison with population values, and it is difficult to separate ethnicity and social deprivation as risk factors for the development of infectious disease. Cases who were not Māori or Pacific cases did not differ significantly in indicators of social deprivation from the total population, yet they might still be more socially deprived than non-Māori and non-Pacific members of the population. Likewise, even though Māori and Pacific cases of SAB were more socially deprived than cases of other ethnic groups, they too may be more socially deprived than the Māori and Pacific population of Auckland. For example, the proportion of Māori and Pacific cases of SAB who lived in areas with NZDep2006 scores of 10 was double that of the total population.

In the current study, ethnic variation in the incidence of SAB was largely due to differences in the incidence of diabetes and renal disease between ethnic groups. However, community-onset SAB comprised a larger proportion of Māori and Pacific cases of SAB, and in turn was associated with a number of indicators of social deprivation. Invasive disease caused by SA is an excellent model on which to base future research regarding the impact of social deprivation on ethnic variation in infectious diseases in New Zealand. Further understanding is a high priority and will lead to a reduction in the incidence of SAB, an illness with a mortality rate of 15%.
Chapter 2: The burden of *S. aureus* skin and soft tissue infection in Auckland and evidence of ethnic variation between subtypes of methicillin-resistant *S. aureus* causing wound infection in Auckland

2.1 Introduction

*Staphylococcus aureus* [SA] is well known for its ability to cause serious, invasive disease. The very first edition of the New Zealand Medical Journal, in 1888, contained a case report of rapidly fulminant endocarditis with widespread embolism, that is reminiscent of the SA disease that is still encountered in our hospitals today. The severe end of the spectrum of SA disease in New Zealand has been reported. The other end of the spectrum, asymptomatic nasal carriage, has also been studied in New Zealand to a limited degree. The burden of less severe, but clinically overt, skin and soft tissue infection remains unquantified.

Methicillin-resistant SA [MRSA] infections were initially associated with acquisition within hospitals. However, community acquired MRSA infections have exploded in most parts of the world this decade; very large increases have occurred in Scandinavia. The incidence of infections caused by MRSA in Auckland rose rapidly in the mid 1990s. In 1997 the prevalence of MRSA in the Auckland community was reported to be 5.9%; but it does not appear to have continued to rise as rapidly this decade.

In the late 1990’s SA bacteraemia [SAB] was more common in Māori and Pacific people; the relative risk for SAB was particularly high in Pacific people. The explanation for this is not known, but household crowding and other indicators of deprivation are likely to be important factors. It is, however, possible that Pacific people are exposed to, or infected by, different strains of SA. This might seem an unlikely proposition until one considers that community-acquired infections caused by the Western Samoan phage pattern MRSA [WSPP MRSA] are more prevalent in Māori and Pacific people living in New Zealand. Between 1995 and early 1998 more than half of the cases of MRSA infection or colonisation at Middlemore hospital were of Pacific ethnicity. While this was certainly a mere reflection of the hospital catchment, more than 90% of these MRSA were WSPP MRSA. Higher rates of WSPP MRSA infections have also been reported among Pacific people living in Queensland, Hawaii and Alaska. Fortunately, strains of WSPP
MRSA are typically susceptible to all non-β-lactam antibiotics. Currently, the WSPP MRSA strain designation is determined by pulsed field gel electrophoresis and the WSPP MRSA is typically non-multi resistant, it universally contains SCC\textit{mec}IVa.

The aim of the current study was to estimate the burden of skin and soft tissue infection caused by SA; to estimate the current prevalence of MRSA in the Auckland community; and to describe the basic demographic features of Aucklanders infected with MRSA. Data from wound swab culture results at Auckland’s only community testing laboratory in 2007 were analysed.
2.2 Materials and Methods

The culture results from all wound and skin swabs processed by Diagnostic Medlab between January 1st and December 31st 2007 were reviewed. Diagnostic Medlab [DML] was the only community laboratory provider for the Auckland region in 2007. Superficial swabs from other sites [e.g. ear swabs] and swabs submitted for MRSA screening were not included in the survey. The identification of SA and antimicrobial susceptibility testing were performed according to CLSI methods and interpretive criteria. In addition to testing the susceptibility to β-lactam antibiotics, susceptibility to ciprofloxacin, erythromycin, cotrimoxazole, tetracycline, fusidic acid and mupirocin were also tested.

Non-multi resistant MRSA [nm-MRSA] were those isolates resistant to β-lactam antibiotics only; and multi-resistant MRSA [mr-MRSA] were those resistant to β-lactams and one or more additional antibiotics.

Duplicate swabs that cultured MRSA taken from the same person were excluded from analysis; in the rare circumstance that the antimicrobial susceptibility profile differed between duplicates, the strain with the greater resistance was included.

Demographic information [age, gender and ethnicity] was obtained for all patients from whom MRSA was isolated and whose National Health Index number [NHI] was available. Ethnicity data from the NHI has been utilised in a number of important publications regarding a range of topics, for example, mortality and cancer rates. The NHI might contain more than one ethnicity for each person, because people may self-report more than one ethnic group; people were prioritised into five ethnic groups, in accordance with Ministry of Health categorisation. These groups were Māori, European, Pacific, Asian and other. The NZ European group contained people who identified as NZ European, European, and/or Pakeha.

Comparative demographic information for the total Auckland population was obtained from the 2006 Census dataset, available from: [http://www.stats.govt.nz/Census/2006CensusHomePage.aspx]. Age data were converted into incidence rates and expressed /100,000/year. The expected number of wound swabs performed in each ethnic group was calculated from each ethnic group’s proportion of the population. To estimate the ethnicity of the people who had a wound swab performed, data were collected from 1000 wound swabs selected at random from those whose NHI number was available.

Tests of statistical significance were performed using SPSS 16.0 [Chicago, USA].

Ethical approval for this study was provided by the Northern Y ethics committee.
2.3 Results

2.3.1 The prevalence and susceptibility of MRSA in Auckland

Just over half [23853/47047; 51\%] of the wound swab cultures performed in 2007 grew SA, and MRSA was found in 8\% [1872/23853]. Once duplicate swabs were excluded, specimens from 1794 people grew MRSA. If this rate of duplicate swabs [78/1872, 4\%] was also found in those whose swabs grew SA then approximately 22,900 people suffered SA skin and soft tissue infection in Auckland, in 2007. This equated to an incidence of 1754/100,000 people-years.

Over half of the MRSA isolates [979/1794, 55\%] were nm-MRSA. Of the MRSA resistant to additional antibiotics, resistance to erythromycin was encountered most often [Table 2.1].

<table>
<thead>
<tr>
<th></th>
<th>ERY</th>
<th>CIP</th>
<th>TET</th>
<th>COT</th>
<th>FUS</th>
<th>MUP</th>
</tr>
</thead>
</table>

Table 2.1. Resistance to six antimicrobials in 815 MRSA\textsuperscript{1} isolated from wound swabs in Auckland in 2007. ERY=erythromycin, CIP=ciprofloxacin, TET=tetracycline, COT=cotrimoxazole, FUS=fusidic acid, MUP=mupirocin.

\textsuperscript{1}A further 979 MRSA isolates were resistant to \beta-lactams only

2.3.2 MRSA subtype and age

The cumulative distribution function of isolates of the ages of people infected with nm-MRSA differed from the cumulative distribution function of the ages of people infected with mr-MRSA [Kolmogorov–Smirnov test, \( P=0.04 \)] [Figure 2.1A]. There was a constant rate of infection of mr-MRSA [Figure 2.1A] until the age of 80 when the rate increased. In contrast, the rate of infection of nm-MRSA was higher in the first few decades of life and decreased with age [illustrated by the decreasingly positive slope in Figure 2.1A]. When the data were adjusted according to the age of the total Auckland population only the sharp increase in mr-MRSA in the elderly [627/100,000 over the age of 80 years] was observed [Figure 2.1B].
Figure 2.1. [A] The cumulative distribution function of nm-MRSA and mr-MRSA by age; [B] the incidence of nm-MRSA and mr-MRSA isolates/100000 by age decade for 1794 MRSA isolates from wound swabs in Auckland in 2007.
2.3.3 The difference in MRSA subtype by ethnic group

855/1000 [86%] of the randomly selected wound swabs were obtained from people who identified with one of four ethnic groups: Māori, NZ European, Pacific and Asian. The proportion of Māori and NZ European people who had a wound swab performed was not different to their proportion of the total population of Auckland. The proportion of Pacific people who had wound swabs performed [239/855, 28%, 95%CI 25-31] was greater than expected, based on the proportion of Pacific people in the total population of Auckland [14%, one sample t test, \(P<0.001\)] [Table 2.2]. Likewise the proportion of Asian people who had a wound swab performed [39/855, 5%, 95%CI 3-6] was less than the proportion of Asian people in the total population [19%, one sample t test, \(P<0.001\)].

<table>
<thead>
<tr>
<th></th>
<th>Māori</th>
<th>Pacific</th>
<th>NZ European</th>
<th>Asian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mr-MRSA isolates ([n=236])</td>
<td>9 [4, 2-7]</td>
<td>36 [15, 11-20]</td>
<td>191 [81, 75-86]^*</td>
<td>0 [0, 0-2]</td>
</tr>
</tbody>
</table>

^1 855/1000 randomly selected wound swabs obtained from Māori, Pacific, NZ European and Asian people
^2 1071/1794 MRSA isolates were from Māori, Pacific, NZ European and Asian people
^3 one sample t test, \(P<0.001\)
^* Proportion of MRSA isolates greater than the proportion who had a wound swab performed, \(\chi^2\) test, \(P<0.001\)

Table 2.2 The number [%, 95% CI] of wound swabs, MRSA isolates, nm-MRSA and mr-MRSA isolates by ethnicity.

In Pacific and NZ European people, the proportion isolates in each ethnic group that were MRSA differed from the proportion of people who had a wound swab performed in each ethnic group. 239/855 [28%] of the wound swabs were performed in Pacific people; however, 460/1071 [43%] MRSA isolates came from Pacific people \(\chi^2=196.8, 3\) d.f., \(P<0.001\). This was solely due to an increase in nm-MRSA isolates; Pacific people were less likely to be infected with mr-MRSA. Māori people followed the same trend, except they were not more likely to have a wound swab performed than expected.

NZ European people had fewer MRSA isolates [411/1071, 38%] than expected [488/855, 57%] and this result was significant even after the Pacific swabs were removed from analysis \(\chi^2=117.8, 2\) d.f., \(P<0.001\). mr-MRSA was more likely to be isolated from NZ European than expected and they were less likely to be infected with nm-MRSA.
2.3.4 The incidence of nm-MRSA and mr-MRSA infection

Table 2.3 shows the incidence /100,000 of having a wound swab performed and infection with MRSA in each ethnic group in 2007. Pacific people had more wound swabs performed in 2007, compared with Māori, NZ European and Asian people. The rates of MRSA, nm-MRSA and mr-MRSA infection in Pacific people was similar to the rates of MRSA, nm-MRSA and mr-MRSA infection in Māori people after adjustment for the number of wound swabs that were collected during 2007. Māori and Pacific people had higher rates of MRSA infection than NZ European people, but infection was caused by nm-MRSA. The highest incidence of mr-MRSA infection was observed in NZ European people. Asian people had low rates of wound swab cultures and low rates of MRSA.

<table>
<thead>
<tr>
<th></th>
<th>Māori</th>
<th>Pacific</th>
<th>NZ European</th>
<th>Asian</th>
</tr>
</thead>
<tbody>
<tr>
<td>wound swabs performed</td>
<td>3053</td>
<td>6319</td>
<td>3286</td>
<td>783</td>
</tr>
<tr>
<td>MRSA isolates</td>
<td>133</td>
<td>259</td>
<td>59</td>
<td>8</td>
</tr>
<tr>
<td>nm-MRSA isolates</td>
<td>126</td>
<td>238</td>
<td>31</td>
<td>8</td>
</tr>
<tr>
<td>mr-MRSA isolates</td>
<td>7</td>
<td>20</td>
<td>27</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.3. The incidence/100,000/year of wound swabs, MRSA isolates, nm-MRSA and mr-MRSA isolates by ethnicity

The highest incidence of either subtype of MRSA was 627/100000 observed in mr-MRSA in NZ European over the age of 80 years. At ages less than 70 years mr-MRSA infections were rare and were not different between NZ European, Māori and Pacific people [Figure 2.2A]. In contrast, infection of nm-MRSA was more common in Pacific people than NZ European at all age groups [Figure 2.2B].
Figure 2.2. The incidence/100,000/year of [A] mR-MRSA and [B] nM-MRSA for Māori, Pacific and NZ European people by age.
2.4 Discussion

The analysis of raw laboratory data indicates that SA causes a huge burden of skin and soft tissue disease in NZ. Approximately two per cent of Auckland’s population had a wound swab culture performed that grew SA. It is certain that not all of these required treatment, but this number probably provides an underestimate of the true burden. Some people with skin and soft tissue infections caused by SA would not have seen a healthcare practitioner; many who did would not have had a swab performed; and some swabs may have failed to grow SA, perhaps due to concomitant antibiotic treatment.

Pacific people were more likely to have a wound swab performed, and Asian people were less likely to have a wound swab performed, compared to their proportions of the Auckland population. Even after taking this into consideration, there were important differences in nm-MRSA or mr-MRSA rates in different ethnic groups. Māori and Pacific people had higher rates of nm-MRSA than NZ European people across almost all age groups; yet NZ European had higher rates of mr-MRSA in old age.

The very low proportion of Asian people that had wound swabs performed goes some way towards explaining their very low prevalence of MRSA infection. The current study was not able to explore the reasons for this, but it is possible that Asian people in Auckland have reduced access to healthcare, or seek alternative healthcare when unwell. It is even possible that Asian people living in Auckland are less likely to develop skin and soft tissue infections, or have lower rates of SA carriage, a state known to predispose to SA infection.

Likewise, one could only speculate about why Pacific people had a higher number of wound swabs performed than expected. Higher rates of infection with nm-MRSA were evident from a young age; this raises the possibility that the incidence of skin and soft tissue infections, such as impetigo, is not equal among the four population groups.

The laboratory data were analysed in the light of demographic data from 1000 randomly selected wound swabs. This provided an estimate of the likelihood of a wound swab being performed in each population, in an attempt to avoid overestimating ethnic differences in MRSA incidence. If the laboratory data were compared directly with the proportion of the Auckland population one might have erroneously concluded that the incidence of MRSA infection in Pacific people was 3.5 times higher than expected, rather than the more accurate estimate, which indicates that MRSA infection was 1.5 times greater than expected in Pacific people.

The results confirm that Pacific people and NZ European are infected with, or exposed to, different strains of MRSA. An unusual definition of nm-MRSA [resistant to β-lactam antibiotics only] was used to identify WSPP MRSA strains [sequence type 30]. WSPP MRSA strains are typically susceptible to all non-β-lactams and are the dominant MRSA strain in NZ; in 2008, over three-quarters WSPP MRSA were isolated at community laboratories. The AK3 strain
[sequence type 5] is also common in the Auckland community, and a proportion of these have lost their typical resistance to mupirocin. Some AK3 strains would have been classified nm-MRSA by the definition that was chosen. The association between advancing age and mr-MRSA in elderly NZ European is likely to reflect healthcare acquisition, particularly in long-term residential care facilities. The association between nm-MRSA and younger age in Pacific people is likely to reflect acquisition in the community.

No research to date has explored the reason why Māori and Pacific people might be infected with different strains of MRSA compared to other ethnic groups. This finding demands further investigation, particularly when the higher incidence of bloodstream infection found in Māori and Pacific people compared to other ethnic groups is taken into account. Some of the increased burden of diseases caused by SA reported in Pacific people could be attributed to exposure to different strains of SA. Infection with a particular strain of SA might even be related to infection in the Pacific rather than acquisition in Auckland.

The current study has raised a lot of questions that warrant further research. The high burden and severity of disease caused by SA make this a high priority. Of major concern is the finding that Pacific people are more likely to develop community-acquired infection with certain strains of MRSA and at a young age. It is fortunate that, for now, these strains of MRSA remain susceptible to all other commonly used anti-staphylococcal antibiotics.
Chapter 3: The prevalence and diversity of MRSA causing wound infection in Samoa

3.1 Introduction

Infections caused by methicillin-resistant *Staphylococcus aureus* [MRSA] have become a global health concern over the past two to three decades. The epidemiology of MRSA has demonstrated marked geographical variation in the prevalence and genotypes of MRSA [46, 47]; and recent reports from many parts of the world indicate that the prevalence and diversity of MRSA continues to increase [48, 49]. Studies of the global epidemiology frequently have not included MRSA obtained from people living in developing nations. The prevalence and genetic variation of MRSA infection in most Pacific Island nations remains unknown. To date, the only study of MRSA in the Pacific was conducted following a report in 2004 of an increased incidence of MRSA infection in Pacific people in Hawaii [50]. This study showed that the majority of MRSA infections in Hawaii were caused by the USA300 MRSA strain [51].

Relatively high rates of MRSA infection have been reported in Pacific people living outside the Pacific Islands region, including in Alaska, Australia, and New Zealand, and have been attributed to infection with the Western Samoan phage pattern MRSA [WSPP MRSA] [36-38, 40]. WSPP MRSA was identified in New Zealand in 1992 [52]. The first and many subsequent isolates of this strain were from people in New Zealand who had some association with Samoa. The origin of this clone has been postulated [53], and it has spread to Europe and recently to South America [49, 54].

Samoa is an independent nation in the Southwest Pacific, with a population of approximately 180,000 people (Samoan Statistics Department, [http://www.spc.int/prism/wstest/index.htm](http://www.spc.int/prism/wstest/index.htm), Figure 1). Samoa has a small and developing economy comprised predominantly of remittances from Samoan people living overseas, agriculture and tourism. The gross domestic product per capita was $US 2987.9 in 2008 ([http://data.un.org](http://data.un.org)). The life expectancy in Samoa is 74.9 years for women and 68.5 years for men, and the infant mortality rate is currently estimated to be 22.3/1,000 live births. The publicly funded National Health Service is based at the national referral hospital in Apia; in addition there are seven district hospitals. A number of outreach and integrated community health services provide primary healthcare services such as clinics and immunizations.
Samoa gained independence from New Zealand in 1962; and a large Samoan population resides in New Zealand. Samoan people make up the largest population of Pacific people living in Auckland – the largest Polynesian city in the world. In New Zealand infections caused by SA whether methicillin susceptible or resistant, are more common in Māori and Pacific people than other ethnicities. Antimicrobial resistance is routinely monitored in New Zealand, through testing of bacterial isolates from specimens taken for diagnostic purposes. In 2008, 9% of S. aureus were methicillin resistant, although this rate varied considerably between different parts of the country, with rates as high as 13-16% in hospitals in the Auckland area.

In Samoa, as in many other developing countries, there is no systematic monitoring of antimicrobial resistance. The microbiology service at the main public hospital in Apia performs antimicrobial susceptibility testing only when specifically requested by the treating doctor. This testing is infrequent and the results are not regularly collated [interview with laboratory director, 2005]. Antimicrobial susceptibility testing requires equipment, resources, and a high level of expertise and quality assurance – requirements that are beyond the means of many laboratories in the developing world. If there is a high rate of antimicrobial resistance in Samoa, this would have a substantial impact on this relatively poor country with relatively high rates of infectious diseases.

3.2 Aims

This study was performed to assess the prevalence of MRSA isolated from wound swabs from people with soft tissue infection living in Samoa.

The MRSA isolates were characterised to determine the genotypes of MRSA that cause infection in Samoa.
3.3 Materials and Methods

3.3.1 Study procedure

In the summer of 2007-8, a study investigator [James Alesana-Slater, University of Otago] collected isolates of *S. aureus* from people with skin and soft tissue infection in towns and villages around Samoa. Participants were approached in hospitals and health centres, or while attending community-based health clinics: nurse visits to schools or other facilities to provide primary healthcare services such as immunization. At hospitals, participants included both inpatients and outpatients. An opportunistic sampling strategy was employed; all consecutive participants, who attended the study site on the day of the investigator’s visit, were recruited. All attendees at the clinics were asked if they had skin lesions and participants with evidence of infection/inflammation were asked to participate. All participants had clinical evidence of skin infection, although this was not always the primary reason for contact with the health services. In hospitals, doctors and nurses were asked to identify patients with skin and soft tissue infections. If there was more than one infected lesion in the same region [e.g. upper body], only the larger lesion was swabbed. In order to sample a broad range of Samoan society a large number of primary care clinics were visited on both Savai’i and the main island Upolu. The main public hospitals in Apia and Savai’i were also visited.

The patient data collected included demographic information [age, sex, occupation and village of residence] and recent exposure to health services [use of antibiotics in the last month and contact with health services in last three months]. Attempts were made to include people of different ages, gender and occupations.

Ethical approval for the study was given by the Lower South Ethics Committee in New Zealand, and the Health Research Council in Samoa. Participants were given an information sheet about the research and a small card explaining, in Samoan, how to take care of wounds. The study investigator explained the purpose of the study in Samoan and the consent form was printed in Samoan and in English.

3.3.2 Microbiological methods

Swabs were taken after the wound was superficially cleaned with sterile water. Swabs were placed into Amies transport media [Fort Richard, Auckland, New Zealand], stored on ice and transported to the laboratory at one of the two main hospitals. All swabs were processed within 24 hours of collection. Swabs were inoculated onto sheep blood agar containing aztreonam [6 mg L⁻¹] and incubated for 16-24 hours. A catalase test was performed on suspicious colonies and catalase-positive colonies were identified using a latex agglutination test [Staphyloslide™ Latex test kit liquid, BD, Mt Wellington, Auckland, NZ]. *S. aureus* isolates were transported on a nutrient agar slope to LabPLUS, Auckland, New Zealand.
In Auckland, antimicrobial susceptibility testing was performed using disk diffusion for penicillin, and agar dilution breakpoint testing for cotrimoxazole, doxycycline, gentamicin, oxacillin and vancomycin, in accordance with Clinical and Laboratory Standards Institute methods and interpretive standards \(^{41,42,56}\). MRSA were also tested by disk diffusion for susceptibility to erythromycin and ciprofloxacin. Additional fusidic acid and mupirocin disk diffusion susceptibility testing was performed by Helen Heffernen at the Institute of Environmental Science and Research [ESR], Wellington, New Zealand, when required to assist with MRSA strain identification.

All MRSA isolates were characterised by multilocus sequence typing [MLST], as described in Part B, Chapter 5 \(^{57}\).

Helen Heffernen [ESR, Wellington] performed spa typing and, if necessary, pulsed-field gel electrophoresis [PFGE] of Smal-digested genomic DNA, to identify MRSA strains. StaphType software version 1.5 [Ridom GmbH, Würzburg, Germany] was used to assign spa types. spa types were compared using the BURP algorithm and by excluding spa types with <5 repeats and setting a maximum cost of 4 between members of a spa group cluster. PFGE banding patterns were analysed with BioNumerics software version 5.1 [Applied Maths, St-Martens-Latem, Belgium] using the Dice coefficient and unweighted-pair group method with arithmetic averages, at settings of 0.5% optimisation and 1.5% position tolerance.

Categorical data is expressed as proportions and 95% confidence intervals [CI], calculated by the modified Wald method. Statistical significance of comparative categorical data was examined using Fisher’s exact test. The \(\chi^2\) test was used to test for significant variation between the observed and expected numbers of people who grew SA and MRSA in each age group.
3.4 Results

Four hundred and twenty-three swabs were obtained from skin lesions in 399 people at eight locations on the two main islands of Samoa [Upolu and Savai’i] and a smaller island, Manono. The eight locations were the main public hospital in Apia [the capital on Upolu], the main public hospital on Savai’i, five district hospitals around the main islands, and a community-based clinic on Manono.

Skin infections had occurred following falls, cuts, dog bites, car accidents, burns and amputations; or spontaneously in the case of carbuncles, boils, and abscesses. Many were described as “po’u” a Samoan term for tropical ulcer. Twenty-four people had swabs taken from two separate lesions. 20/399 (5%) represented infections of surgical sites, which were hospital acquired; 335/399 (84%) represented community-acquired infections such as impetigo, po’u, boils or infections of wounds resulting from dog bites, machete cuts, or assaults; the remaining 44/399 (11%) cases were not able to be classified further.

*S. aureus* was isolated from 196 of the 423 wound swabs [46%, 95% CI 42-51] and from 187 of the 399 study participants [47%, 95% CI 42-52]. *S. aureus* was more likely to be obtained from children under the age of 5 years [37/53, 70% compared with 150/346, 43%; \( \chi^2 \) test \( p<0.05 \); but was less likely in those currently or recently treated with antibiotics [84/224, 38% compared with 103/175, 59%; Fisher’s exact test \( p<0.001 \)] [Table 3.1].

Thirty-four of the 196 *S. aureus* isolated were methicillin resistant [17%, 95% CI 13-23]. MRSA was isolated from 8% [95% CI 6-11] of the 423 wound swabs and 9% [95% CI 6-12] of the 399 study participants. Almost one in five of the 187 participants with *S. aureus* had MRSA [34/187, 18%, 95% CI 13-24]. There were no important differences in the basic demographic features of people with MRSA, when compared to those with MSSA. Recent antibiotic use and recent healthcare was not associated with an increase in MRSA infections. For example, the proportion of people with recent healthcare exposure who grew MRSA [23/262, 9%] was not significantly different to the proportion of people without recent healthcare exposure [11/137, 8%] [Table 1].

Participants lived in 165 villages around Samoa. *S aureus* was detected in participants from 102 villages. In terms of place of domicile, MRSA was widespread throughout Samoa and was isolated from people from 27 villages. There did not appear to be any geographical clustering associated with residence close to the main public hospital in Apia, although MRSA was more likely to be isolated from people from Upolu, the main island on which Apia is located, than from residents of Savai’i, the less populated and less developed island. MRSA was isolated from 11% of people from Upolu, but from only 3% of people from Savai’i.
### Table 3.1. The demographic features of study participants and prevalence of methicillin-susceptible and MRSA in Samoa.

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>Total study population</th>
<th>SA(^1) positive</th>
<th>MSSA(^2)</th>
<th>MRSA(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>399</td>
<td>187 [47, 42-52]</td>
<td>153 [38, 34-43]</td>
<td>34 [9, 6-12]</td>
</tr>
<tr>
<td>Age &lt;5 years</td>
<td>53</td>
<td>37 [70, 56-81](^4)</td>
<td>33 [62, 49-74]</td>
<td>4 [8, 3-18]</td>
</tr>
<tr>
<td>School age</td>
<td>93</td>
<td>50 [54, 44-64]</td>
<td>44 [47, 38-57]</td>
<td>6 [7, 3-14]</td>
</tr>
<tr>
<td>&gt;60 years</td>
<td>58</td>
<td>21 [36, 25-49]</td>
<td>16 [28, 18-40]</td>
<td>5 [9, 3-19]</td>
</tr>
<tr>
<td>Antibiotic treatment in the previous month</td>
<td>224</td>
<td>84 [38, 31-44](^4)</td>
<td>67 [30, 24-36]</td>
<td>17 [8, 5-12]</td>
</tr>
<tr>
<td>Healthcare contact in the previous 3 months</td>
<td>262</td>
<td>108 [41, 35-47]</td>
<td>85 [32, 27-38]</td>
<td>23 [9, 6-13]</td>
</tr>
</tbody>
</table>

\(^1\) SA, S. aureus  
\(^2\) MSSA, methicillin-susceptible S. aureus  
\(^3\) MRSA, methicillin-resistant S. aureus  
\(^4\) \(\chi^2\) test, \(P<0.05\) S. aureus positive versus total study population  
\(^5\) Fisher exact test, \(P<0.05\) S. aureus positive versus total study population

Among the 153 MSSA isolated from 187 participants, 124 [81\%] were resistant to penicillin, but resistance to any other antibiotic was uncommon. Among the 34 MRSA isolates, 22 were resistant only to β-lactams. Among the remaining 12 MRSA, nine were ciprofloxacin and erythromycin resistant, two were erythromycin resistant and one was ciprofloxacin resistant. All MRSA were susceptible to cotrimoxazole, doxycycline, gentamicin and vancomycin.

The strains identified among the 34 MRSA isolates are shown in Table 3.2. The majority of the MRSA [25/34, 74\%] were internationally recognised strains, with the exception of the MLST type ST1 MRSA. Seven of the nine ST1 isolates had indistinguishable PFGE profiles. The ST 1 MRSA isolates had 82\% homology by PFGE typing with another ST1 MRSA strain common in the region - the WR/AK1 New Zealand MRSA [Australian WA MRSA-1]. There was an association between antimicrobial resistance profile and strain [Table 3.2]. All isolates of the USA300 strain were ciprofloxacin resistant and most were also erythromycin resistant. All of the Queensland clone and Southwest Pacific clone MRSA, and eight of the nine ST1 MRSA isolates, were resistant only to β-lactams.
<table>
<thead>
<tr>
<th>Number [%] of isolates</th>
<th>Strain</th>
<th>MLST type</th>
<th>spa type</th>
<th>Antimicrobial susceptibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 [29]</td>
<td>USA300</td>
<td>ST8</td>
<td>t008 [10]</td>
<td>9 resistant to ciprofloxacin and erythromycin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 resistant to ciprofloxacin</td>
</tr>
<tr>
<td>9 [26]</td>
<td>Queensland clone</td>
<td>ST93</td>
<td>t3949 [5] t202 [4](^2)</td>
<td>all resistant only to β-lactams</td>
</tr>
<tr>
<td>9 [26]</td>
<td></td>
<td>ST1</td>
<td>t1853 [7] t6080 [1](^4)</td>
<td>8 resistant only to β-lactams 1 resistant to erythromycin</td>
</tr>
<tr>
<td>4 [12]</td>
<td>WSPP/Oceania clone</td>
<td>ST30</td>
<td>t019 [4]</td>
<td>all resistant only to β-lactams</td>
</tr>
<tr>
<td>2 [6]</td>
<td>AK3</td>
<td>ST5</td>
<td>t002 [1] t1265 [1](^5)</td>
<td>1 resistant to erythromycin 1 resistant only to β-lactams</td>
</tr>
</tbody>
</table>

\(^1\) International MRSA strain designations, except for AK3 which is a designation for a community-associated MRSA strain common in New Zealand.

\(^2\) spa type t3949 [11-17-23-17-17-17-16-16-25] is a single-repeat variant of t202 [11-17-23-17-17-16-16-25]. The extra repeat is shown in boldface.

\(^3\) Only eight of the nine ST1 MRSA were available for spa and PFGE typing. Seven of the eight available isolates had indistinguishable PFGE profiles and the eighth shared 87% homology.

\(^4\) spa type t6080 [07-23-21-17-13-34-16-13-33-13] is a single-repeat variant of t1853 [07-23-21-17-13-34-16-13-33-13]. The extra repeat is shown in boldface.

\(^5\) spa type t1265 [26-23-17-13-17-17-12-17-17-12-16] is a variant of t002 [26-23-17-17-17-17-12-17-17-12-16]. The differing repeats are shown in boldface.

Table 3.2. Strains identified among the MRSA isolated in Samoa

In New Zealand, in 2007 six different MRSA strains caused the majority of infections and WSPP MRSA was the most common MRSA strain \(^3\). Figure 3.1 shows the proportion of predominant MRSA STs in Samoa and in New Zealand in 2007. It was expected that the MRSA types in Samoa would be similar between the two countries; however, this was not the case. In particular USA300, ST8 MRSA caused almost one third of MRSA infections in Samoa, yet in New Zealand USA300, ST8 MRSA infections were rare.
Figure 3.1 The proportion of predominant MRSA STs in Samoa (current study) and New Zealand in 2007. The PFGE pattern of ST1 strains in Samoa differed from the PFGE pattern of ST1 strains in New Zealand.
3.5 Discussion

There is considerable concern internationally about rising levels of antimicrobial resistance. Consequences include treatment failure, the need to use newer antibiotics to achieve treatment goals, the adverse effects frequently associated with these newer drugs, increased expenditure, and longer hospital stays.

In developing countries like Samoa there is the added concern that, when resistance to cheaper antibiotics becomes widespread, the more expensive antibiotics are simply not available. As in many other developing countries, consumption of antibiotics is extremely high in Samoa. More than 60% of all prescriptions dispensed in hospitals and private pharmacies include an antibiotic. There are no current data available on the prevalence of antibiotic resistance in Samoa. In other settings high levels of antibiotic use correlate with high levels of resistance. In addition, inadequate dosing and poor adherence by patients may also increase the development of resistance. There is anecdotal evidence that these practices are also common in Samoa.

There is little data available on prevalence of MRSA in the community in developing countries. The World Health Organization has identified a need to strengthen monitoring of prescribing and resistance and has funded pilot programmes to achieve this. In this study in Samoa, MRSA were isolated from 9% of all participants and 18% of participants from whom SA was isolated had a MRSA. The age groups of the participants indicate that a broad range of people were included in the study. Carrying out such research is difficult in a developing country such as Samoa, where research and laboratory infrastructure are not ideal. The wound swabs were collected by a Samoan student at the University of Otago, who returned to Samoa for the summer holiday. Ability to speak Samoan, knowledge of Samoan culture and previous experience of living in Samoa were essential for the success of the project. All of the antimicrobial susceptibility testing and MLST genotyping was performed by the author of this thesis, who also performed all of the analysis. The MLST analysis adequately captured the diversity of MRSA in Samoa, but PFGE and spa typing, performed by Helen Heffernen, ESR, matched the MRSA strains with internationally recognised designations.

The MRSA in Samoa were surprisingly diverse. The diversity of MRSA that caused skin and soft tissue infection in Samoa was similar to that found in Denmark, a country of over 5 million people in central Europe. It was expected that WSPP MRSA would be the predominant MRSA strain in Samoa, but these were in the minority. The high prevalence of the Queensland clone MRSA might be a reflection of the amount of travel between Samoa and Australia, where the Queensland clone is now common in the states of New South Wales and the Northern Territory in addition to Queensland. The USA300, ST8 MRSA strain is also common in Samoa and is likely to have entered via Samoan people travelling back and forth from the West coast of USA. There is a moderately large population of Samoan people living in Los Angeles and people living in American Samoa frequently travel to the USA for healthcare; anecdotally, travel between Samoa...
and American Samoa is common. However, the large number of isolates of the Samoan ST1 MRSA clone and the isolation of almost equal numbers of three different MRSA clones suggest that the situation may not be that simple.

The Samoan ST1 MRSA isolates were distinct by both PFGE and spa typing from the community-associated ST1 MRSA strain commonly found in Australia and New Zealand, WA MRSA-1. The Samoan ST1 MRSA isolates shared only 82% homology by PFGE with WA MRSA-1. WA MRSA-1 isolates are typically spa type t127, which did not cluster by BURP analysis with the spa types [t1853 and t6080] of the Samoan ST1 MRSA. In addition, the WA MRSA-1 strain is characterized by fusidic acid resistance and often also mupirocin and/or erythromycin resistance, whereas the Samoan ST1 MRSA isolates were mainly resistant only to β-lactams. The Samoan ST1 MRSA isolates were also distinct from the USA400 ST1 MRSA strain with which they shared only 80% homology by PFGE. It is possible that the Samoan ST1 MRSA strain has its origins in the Pacific; MRSA with spa type t1853 have also been isolated in New Zealand, mainly from patients in the Auckland area, since at least 2008.

Reducing the burden of resistant organisms is challenging but is of vital importance. The mortality rate of serious infections caused by MRSA is greater than those caused by MSSA. Furthermore, MRSA infections carry a far greater economic cost in developed countries. There are no similar studies investigating these consequences in the developing world, although the burden of disease is likely to be considerable. Reliable surveillance data is essential to monitor, control and manage antimicrobial resistance. The two interventions most likely to reduce the prevalence of MRSA are improved prescribing and infection control; both are reliant on reliable surveillance data.

Thus, there is a clear need to assist developing countries to perform quality antimicrobial susceptibility testing and surveillance. At the very least, one in five of the participants in our study might have benefited from more appropriate prescribing for their MRSA infection. Unfortunately, the susceptibility testing was performed in Auckland and the results were not available in time to be of service to the study participants. However, on the positive side, the majority of MRSA in Samoa appear not to be multi-resistant, with the exception of isolates of the USA300 strain which were resistant to ciprofloxacin and erythromycin.

This study will provide a starting point for future research into antimicrobial resistance in the Pacific, and provide impetus for initiatives to improve antibiotic use in Pacific Island nations. Antimicrobial resistance is a global problem, that does not respect national boundaries, and consequently it is imperative that countries assist each other to address the problem. This study has identified a surprisingly wide range of genotypes of MRSA causing wound infections in a small Pacific Island nation that differed from the genotypes of MRSA that cause infections in New Zealand. These differences might reflect substantial differences between the total SA populations in Auckland and Samoa.
Chapter 4: The prevalence and demographic features of S. aureus carriage in Auckland

4.1 Introduction

*Staphylococcus aureus* [SA] causes many serious infections in New Zealand every year. SA is a common cause of skin and soft tissue infection, bone and joint infection, endocarditis and healthcare associated infection. Infection with SA is much more common than the incidence of SAB suggests and represents a spectrum of disease from asymptomatic colonisation [SA carriage] through skin and soft tissue infection to invasive disease which may result in SAB. Asymptomatic nasal colonisation [SA carriage] is a risk factor for the development of SA disease. Nosocomial SAB is more common in carriers 71-73; and there is evidence that treatment to resolve colonisation can reduce the incidence of subsequent surgical wound infection 74, 75. Thus, a simple model of infection with SA is one of exposure via direct or indirect contact followed by a period of asymptomatic colonisation prior to contamination of damaged tissue and/or a medical device leading to the development of disease. SAB represents only a very small proportion of SA infection, approximately one in one thousand people colonised with SA develop SAB.

SA colonisation can be defined as the presence of SA at any of a number of anatomical sites in healthy and asymptomatic people. The rates of colonisation vary with study methodology and the population under study; but the prevalence of nasal colonisation is generally between 20% and 37% of the healthy population 76-78. Most studies of SA carriage have involved nasal carriers; and the anterior nares are the site most frequently colonised by SA 78. A recent study, conducted in hospital inpatients and hospital staff, which suggested that throat colonisation was more common than nasal colonisation was not confirmed by a subsequent study 79, 80; however, a small proportion of SA carriers have throat colonisation without concomitant nasal colonisation. Studies will continue to focus on SA nasal carriage because important immunological differences have been identified between persistent nasal carriers and persistent nasal non-carriers.

Persistent carriers of SA appear to be infected with a single clone of SA. Hu et al, found that SA strains isolated from persistent carriers over a two-year period did not differ in their genotype 81. This was also the case in other longitudinal studies of colonisation 82, 83. In a recent report of a study conducted in Malaysian university students, no persistent carriers were identified 84. Twelve students were carriers of SA on two occasions a month apart; however, in every case the paired
strains of SA isolated from each individual were different. This study demonstrated that colonisation with SA in intermittent carriers was surprisingly dynamic, and that some people are susceptible to sequential infection with different SA strains.

4.1.1 Immune and genetic differences between persistent carriers and persistent non-carriers

Immune responses to SA are different in persistent carriers and persistent non-carriers. Nasal colonisation does represent true infection and is associated with a host immune response. Cole et al screened healthy people in California, and found that carriage of SA was more common in white people than in Hispanic people and African Americans\(^85\). The investigators found evidence of a greater neutrophil response in nasal secretions obtained from carriers of SA compared with non-carriers. Despite this finding, the nasal secretions from carriers of SA were less toxic to SA in vitro.

A study of experimental inoculation of SA into the nares of healthy volunteers found that nasal infection persisted longer in people who were susceptible to the smallest inoculum\(^86\). Participants in this study were less susceptible to repeat inoculation with the same strain of SA, presumably as a consequence of adaptive immunity, yet remained susceptible to infection with other strains. This result was confirmed recently\(^87\); but elevated levels of antibodies against staphylococcal toxins or adhesion molecules that might protect against colonisation were not found in persistent non-carriers\(^88\). For example, antibody concentrations against clumping factor B [ClfB] were higher in persistent carriers than in persistent non-carriers.

The finding of an association between polymorphisms in the interleukin-4 gene and nasal carriage provided further evidence that host immunity is an important factor in nasal carriage of SA\(^89\). Polymorphisms of the glucocorticoid receptor gene, the mannose-binding lectin gene, and the serine protease inhibitor of complement factor 1 gene [C1INH] are also associated with an increased risk of nasal carriage of SA\(^90-92\). It is not known whether polymorphisms in these genes are directly involved in susceptibility to nasal infection or whether they are markers for other, yet to be discovered, genetic differences.

4.1.2 Bacterial determinants of nasal colonisation

Quinn and Cole highlighted the role that the bacterium plays in nasal colonisation\(^93,94\). They used an in vitro model of nasal epithelial cell culture to compare a “carrier strain” of SA with a strain of SA that did not cause experimental colonisation and a strain of \(S.\) epidermidis. They found that the carrier strain displayed better attachment and growth on the epithelial cells and that there was a delay in the up-regulation of toll-like receptor-2 and a reduction in the expression of IL-1 and defensins by epithelial cells infected with the carrier strain. When the epithelial cells were pre-treated with interleukin-1\(\beta\) the growth of the carrier strain was reduced by a greater degree than the other staphylococcal strains. There were a number of differences between the genomes of the
carrier and non-carrier strains. In particular a putative pathogenicity island was only found in carriage strains.

Despite the finding that antibody against ClfB did not protect against colonisation, ClfB has an important role in nasal colonisation. ClfB is a surface molecule that binds to human-type I cytokeratin 10, present on the surface of epithelial cells. Wertheim et al found that a ClfB knockout SA mutant was less able to persist after inoculation into the nares of healthy volunteers than the wild type strain.

4.1.3 Demographic variation in nasal colonisation

Variation in the overall prevalence of colonisation or in the prevalence of colonisation by particular strains of SA could contribute to the ethnic variation in the incidence of invasive disease in New Zealand. Ethnic variation in SA carriage has not been studied in New Zealand. In Queensland, Australia, Aboriginal people had lower rates of colonisation than European people; and the prevalence of SA carriage was lower in people from a socially deprived area of Queensland compared to the prevalence in people living in an affluent area. Screening for nasal carriage of SA was performed on a large national sample of over 9,000 non-institutionalised individuals in the USA. African Americans were less likely than Hispanic or white Americans to be carriers of SA, and deprivation was not associated with carriage of SA.

The current study of nasal carriage of SA was performed in healthy population volunteers to investigate the rates of carriage and to explore demographic differences between SA carriers and non-carriers in New Zealand. The strains of SA isolated in the current study, were analysed in a further study of the genetic population structure of SA in Auckland [Part B, Chapter 5].
4.2 Materials and Methods

Ethical approval was provided by the Northern Y ethics committee of the New Zealand Ministry of Health. The University of Auckland ethics committee deferred to the Northern Y ethics committee and granted access to University grounds for recruitment purposes.

4.2.1 Study participants

In order to obtain a sample of SA carriers with similar demographic characteristics to the cases of SAB recruited in the major study [Chapter 1], participants were recruited in public places in Central and South Auckland. The investigator visited public sites several days a week; the most commonly visited sites were Aotea square in central Auckland, the University of Auckland grounds and Albert Park, Mission Bay, public spaces between Manukau City mall and council offices, Hunter’s corner in Papatoetoe and sports grounds in Mangere East. All sites were visited several times between April 2007 and January 2008. Participants who had hospital contact, which might have influenced the prevalence of carriage, were excluded from the study.

After informed consent was obtained, a nasal swab was performed. One swab was inserted into both nostrils, being careful to sample the mucosa on the nasal septum adjacent to the nasal ostium, the preferred habitat of SA. The principal investigator and a medical student [Miss Nicola Best, University of Auckland] performed all study procedures. The consent process and all study procedures were performed at the point of contact. Participants completed a brief questionnaire about demographic factors, home address, household size, access to healthcare, and prior SA infection.

4.2.2 Culture of SA

Nasal swab specimens were transported at ambient temperature in sterile Amie’s media [Fort Richard, Otahuhu, Auckland, NZ] and were swabbed onto mannitol salt agar [MSA] [Fort Richard] within four hours of collection. Following overnight incubation at 37°C presumptive SA colonies were identified using the Staphyloslide™ Latex test kit [BD, Mt Wellington, Auckland, NZ]. MSA plates that did not display growth of SA were incubated for a further 24 hours. The identity of positive isolates was confirmed by the production of coagulase and staphylococcal DNase.

Antimicrobial susceptibility was tested using a combination of agar breakpoint dilution [oxacillin, erythromycin, cotrimoxazole, doxycycline and gentamicin] and disc diffusion methods [penicillin] according to Clinical and Laboratory Standards Institute performance and interpretive criteria.

52
4.2.3 Analysis

Information from the study population was compared with population data from the 2006 census. Ethnicity was self-reported by participants who selected their ethnic group[s] from a list; the same list used in the 2006 New Zealand Census. Ethnicity was prioritised using the method advised by the Ministry of Health. To allow comparison with 2006 census data, the ethnicity data collected was grouped according to level 1 classification. The NZ European ethnic group also contained European and Pakeha.

The NZ index of deprivation (NZDep2006), an ordinal score of 1 to 10 derived for small residential blocks using data obtained from the 2006 Census, was used to estimate the deprivation score of the neighbourhood of each participant [see Methods Chapter 1]. The home address of each participant was used to determine their residential block and the NZDep2006 score for each residential block was obtained from the published values.

A bedroom occupancy rate was calculated from the ratio of the number of household occupants resident on at least three nights per week to the number of bedrooms.

Statistical analysis was performed using SPSS 16.0 [Chicago, USA]. 95% confidence intervals were calculated using the modified Wald method [Graphpad online calculator http://www.graphpad.com/quickcalcs/ConfInterval2.cfm]. Fisher’s exact test was used to test the statistical significance of categorical data, and the Wilcoxon signed rank test was used to compare non-parametric data.
4.3 Results

Approximately half of the people approached dismissed conversation with the investigator immediately; but all of the people who did stop to listen to the investigator agreed to take part in the study. 424 subjects participated in the study, their median age was 22 [interquartile range 19-32] and 255/426 [60%] were female. When compared to the total population, people between the ages of 15 and 34 years were over represented in the sample, and there were no young children included in the study [Figure 4.1].

Figure 4.1. The age distribution of the 424 study/100,000 population members and the proportion of participants who were SA carriers in each age group.

The ethnicity of the 424 participants is shown in Table 4.1. Māori, Pacific and NZ European people were under-represented in the sample when compared to their proportions in the total population of Auckland; Asian people were over-represented [one sample T test P<0.01] [Table 4.1].
<table>
<thead>
<tr>
<th>ethnic group</th>
<th>number [%], 95% CI</th>
<th>proportion of the total population¹</th>
<th>SA carriage [%], 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori</td>
<td>29 [7%, 5-10]</td>
<td>11%</td>
<td>6 [21%,10-39]</td>
</tr>
<tr>
<td>Pacific</td>
<td>44 [10%, 8-14]</td>
<td>14%</td>
<td>8 [18%,9-32]</td>
</tr>
<tr>
<td>NZ European</td>
<td>159 [38%, 33-42]</td>
<td>61%</td>
<td>33 [21%,15-28]</td>
</tr>
<tr>
<td>Asian</td>
<td>146 [34%, 30-39]</td>
<td>13%</td>
<td>26 [18%, 12-25]</td>
</tr>
<tr>
<td>Other</td>
<td>46 [11%, 8-14]</td>
<td>1%</td>
<td>5 [11%, 4-24]</td>
</tr>
</tbody>
</table>

¹The proportion of each ethnic group in the sample differed from their proportion in the total population, one sample T test, P<0.01.

Table 4.1. The proportion of SA carriers by ethnicity in 424 participants; and the proportion of each ethnic group in the total population of Auckland.

### 4.3.1 SA nasal carriage

78/424 [18%; 95% CI, 15-22] SA carriers were identified. The median age of SA carriers [22 years, interquartile range 19-33] was not different from the median age of non-carriers [22 years, interquartile range 19-32]. There were no significant differences in the rates of SA carriage between different age groups [Figure 4.1].

35/78 [45%, 34-56] of SA carriers were female, less than the proportion of females in the study population [258/424, 61%, 56-65, Fisher’s exact test, P=0.002]. The proportion of SA carriers in each ethnic group was not different between ethnic groups.

Only 1/78 [1%] of the SA isolates was methicillin resistant [MRSA]. This strain was isolated from a 32 year-old Indian male, who had lived in New Zealand for 7 years, and who did not report any healthcare contact.

### 4.3.2 Deprivation and household crowding

The proportion of SA carriers who held a current community services card [25/78, 32%, 23-43] was not different to the proportion of non-carriers [124/346, 36%, 31-41]. Likewise the proportion of participants who held private health insurance was not different between SA carriers [29/78, 37%, 27-48] and non-carriers [134/346, 39%, 34-44]; neither was the proportion of participants who were registered with a general practice [SA carriers, 60/78, 77%, 66-85; non-carriers, 251/346, 73%, 68-77].

396/424 [93%] participants provided their home address, and the NZDep2006 scores of their neighbourhoods were able to be determined. The median deprivation score of the SA carriers [5, interquartile range 3-8] was lower than the non-carriers [6, interquartile range 3-8, Wilcoxon signed ranks test, Z=-7.154, P<0.01]. This difference did not persist when the deprivation scores were placed into three groups from least deprived [NZDep2006 score 1-3] to the most deprived [NZDep2006 score 7-10] [Figure 4.2].
The mean bedroom occupancy rate did not differ between SA carriers [1.14, s.d. 0.41] and non-carriers [1.15, s.d. 0.39].

![Figure 4.2](image)

Figure 4.2 The deprivation category of SA carriers [n=70] and non-carriers [n=326].

<table>
<thead>
<tr>
<th>Deprivation category</th>
<th>SA carrier</th>
<th>SA non-carrier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40%</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>3</td>
<td>35%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Category 1 contains NZDep2006 scores of 1, 2 and 3; category 2 contains scores of 4, 5 and 6; category 3 contains scores of 7, 8, 9 and 10.

4.3.3 Recent healthcare exposure

Recent healthcare contact was not associated with an increased risk of SA carriage. There was no significant difference in the proportion of SA carriers who had been admitted to hospital in the previous 12 months [3/78, 4%, 1-11] compared to non-carriers [33/346, 10%, 7-13]; and there was no significant difference in the proportion of people who reported a prior infection caused by SA [SA carriers, 3/78, 4%, 1-11; non-carriers 11/346, 3%, 2-6]. 12/78 [15%, 19-25] SA carriers had seen their general practitioner in the previous month compared with 92/346 [27%, 22-32] non-carriers [Fisher’s exact test, $P=0.04$].
4.4 Discussion

No previous studies of SA carriage in the healthy population in New Zealand have been reported. Based on the results of the current study, over 230,000 healthy population members in Auckland are currently carriers of SA. The prevalence of SA carriage in Auckland, 18%, was at the lower end of the range of prevalence rates found in other populations. A large study conducted in the USA in 2006, found the prevalence of nasal carriage to be 32%; and in Queensland in 2005/2006 the prevalence was 28%. The lower rate found in the current study was not likely to be a result of the nasal swab sampling strategy. The recovery of SA from nasal swab specimens used the same microbiological method of similar studies; and a medical practitioner performed all of the nasal swabs. It is also unlikely that the sampling strategy provided estimates of the prevalence of SA carriage that were not representative of the total population.

The prevalence of SA carriage did not differ between different age groups in the current study, but age is an important determinant of SA carriage. The highest prevalence of SA carriage occurs in the first few months of life but then declines over 6-12 months to rates that are similar to the adult population. Despite an increase in the incidence of disease caused by SA in old age, SA carriage is less common in adults over the age of 60. Thus, the age distribution of the sample in the current study is likely to provide a reliable estimate of the prevalence of SA carriage in the population. In the current study, females were less likely to be colonised with SA than males. This finding is consistent with other international studies; even in the first year of life, males were more likely to be SA carriers than females. The reason for the difference in colonisation between males and females is not known, but this difference might contribute to the lower incidence of SAB reported in females.

The current study did not show any relationship between SA carriage and holding a Community Services card [CSC], having private health insurance or the bedroom occupancy rate. When the home addresses of the participants were used to estimate the NZDep2006 scores of each participant's neighbourhood, the level of deprivation was similar for SA carriers and non-carriers. Social deprivation does not appear to play an important part in the development of SA carriage in Auckland. This finding was consistent with a previous study in Queensland, which found lower rates of SA colonisation in people living in socially deprived areas compared to affluent areas. In India, indicators of deprivation such as family income, household crowding and residence in a slum were not different between SA carriers and non-carriers. In a large cohort of Dutch infants, maternal education level, cigarette smoking, and breastfeeding did not influence the development of nasal carriage in the first year of life. Carriage of Neisseria meningitidis in household contacts of people with meningococcal disease in Auckland was more common in smokers and Māori and Pacific people, but was not otherwise associated with deprivation.

Only one carrier of MRSA [1/424, 0.2%] was identified in the current study; the prevalence of MRSA carriage was lower than the prevalence of clinically significant MRSA infections. In 2007,
8% of the community laboratory isolates of SA were MRSA [Chapter 2]. Contemporary studies have also found low rates of MRSA colonisation; in Queensland in 2005, the prevalence of MRSA colonisation was 0.7% ⁹⁷; in the USA between 2000 and 2002, the prevalence of MRSA carriage was 0.8% ⁹⁹. It is likely that the rate of MRSA carriage is increasing in parallel with community-acquired MRSA infections, yet the disparity between the prevalence of MRSA carriage and the frequency of MRSA isolation from clinical specimens suggests that diseases caused by SA may not be a simple reflection of the prevalence of SA carriage. It is possible that MRSA strains are more virulent and more likely to progress from carriage to disease. Comparison of SA genomes might identify further differences between strains that cause nasal colonisation and other diseases ⁹⁵.

The current study has several strengths and weaknesses. A large number of people with diverse demographic characteristics were included in the sample, and the sample almost entirely consisted of healthy population members. As a result, the study population demographics were not well matched with the total population of Auckland. For example, no young children and very few people over the age of 65 years were included in the study. Most of the study participants were young adults between the ages of 15 and 34 [Figure 4.1]. However, this age group is likely to provide estimates of the population prevalence of SA carriage that were close to maximal, with the exception of young infants.

The other major difference between the study sample and the total population of Auckland relate to the ethnicity of the participants. New Zealand European people were significantly under-represented and Asian people were over-represented. It is possible that the low proportion of NZ European people in the sample contributed to lower rates of SA carriage than expected. In Queensland, the prevalence of SA carriage was highest in Europeans [32%] and lowest in Aboriginal people [18%] ⁹⁷. It is unlikely that the high proportion of Asian people in the sample contributed to the low prevalence; studies from India [52%] and China [33% and 18%] have found prevalence rates of SA carriage similar to, or higher than, the prevalence in the current study ¹⁰⁴, ¹⁰⁶, ¹⁰⁷.

The current study did not provide an explanation for the ethnic variation in invasive SA disease in Auckland. The simple model of SA disease arising from the huge number of people who are colonised with SA is not consistent with the lack of differences between SA carriers and non-carriers identified in this study. The lack of ethnic variation and variation by deprivation in SA carriage is not consistent with ethnic variation in the prevalence of SAB. Furthermore, the prevalence of MRSA carriage is lower than the rate predicted from the prevalence of MRSA skin and soft tissue infection in Auckland. Future research should to examine the reasons why females are less likely to be SA carriers than males; in addition, improved understanding of the differences between the SA strains that cause nasal carriage and those that cause disease is crucial to determining how an organism that predominantly causes asymptomatic colonisation causes serious illness.
Part B: Population genetics of 

*Staphylococcus aureus*

Chapter 5: The genetic structure of *Staphylococcus aureus* populations from the Pacific and comparison with the rest of the world

5.1 Introduction

The main aims of this research were to determine whether ethnic variation in SA bacteraemia [SAB] is caused by differences in the strains of SA responsible for infection; to look for evidence of evolution and adaptation of SA in Auckland and the wider Pacific context, and to determine the processes that shape the evolution of SA. Similar analyses have been performed in other locations, but the sample of SA analysed in the current study is unique and has been collected carefully in conjunction with clinical and demographic information regarding the host.

The SA strains collected for the current study were genotyped by multi-locus sequence typing [MLST]. MLST is a powerful tool for grouping and categorising SA strains, but it is more than simply a method of strain typing. The nucleotide sequence that results from MLST provides a direct measure of allelic variation and this allows data to be interpreted using population genetic theory; thus providing insight into the evolutionary forces [mutation, recombination, migration, selection and drift] shaping SA populations.

Elucidation of the genetic structure of diverse of bacterial species has flourished in recent times. Sequencing technology has improved and become more accessible. Furthermore, increased computing power has allowed the development of sophisticated analytical methods. This Chapter describes the genetic structure of SA populations in Auckland and the Pacific Island nations of Fiji, Samoa and Tonga. The genetic structure of SA is examined on a number of levels: within the host [nasal swab and blood culture isolates]; within the local population of Auckland; within the Pacific region and finally within the global population of SA.

This introduction will briefly review the prior use of MLST to study SA populations, the global SA population, and the analytical methods used in this Chapter, together with the population genetic theory relevant to each method.
5.1.1 Multi-locus sequence typing [MLST]: comparison of invasive disease isolates and nasal carriage isolates in Oxford

MLST was developed as an extension to multi-locus enzyme electrophoresis [MLEE] and was first used to compare isolates of *Neisseria meningitidis* [109]. MLST had the advantage over MLEE that it could detect much greater variation between isolates, it was easier to perform and the results could be easily compared between research centres. Enright et al then used the SA MLST method to investigate 155 isolates of SA from patients with invasive disease [57]. Seven housekeeping genes, selected to represent the core genome, were sequenced and each unique sequence was given an allele number. The allele numbers for each of the seven MLST genes were used to derive the sequence type [ST] and STs that had five or six alleles in common were grouped into clonal complexes [CC].

Enright and colleagues then compared isolates responsible for community-acquired invasive disease with isolates responsible for hospital-acquired invasive disease and isolates obtained from asymptomatic nasal carriers [110-112]. The entire collection of strains from hospital inpatients and healthy blood donors was performed in Oxford, England. The total sample of 334 SA isolates was found to comprise 75 distinct STs which were grouped into 11 CCs on the basis of shared ancestry [112]. The invasive disease isolates and nasal carriage isolates were equally distributed in each CC, indicating that there was no difference between the “core” genomes of SA that caused invasive disease or nasal colonisation in Oxford. Analysis of closely related STs showed that the evolution of SA was largely driven by point mutation within clonal complexes. This result was confirmed by comparing phylogenetic trees for each gene locus, which demonstrated that the alleles of related STs generally grouped together on each gene tree. The authors estimated that the evolution of new alleles by point mutation was 15 times more common than by recombination.

Subsequently the same collection of strains, from asymptomatic colonised people and people with invasive disease, was examined for the presence of genes encoding 33 virulence factors [113]. Genes for eight virulence factors were more common in the isolates that caused invasive disease, and the presence of genes for seven of these virulence factors remained significant after stratification of the isolates based on their STs. Overall, the proportion of strains that caused invasive disease rather than nasal colonisation increased with an increase in the number of virulence factors present in the strains. This study indicated that there were differences in the “accessory” genome between the strains that caused invasive disease and those responsible for asymptomatic nasal colonisation. The study also demonstrated that horizontal transfer of virulence genes, many of which are chromosomal, between different STs was a common event.

Further analysis was performed on a subset of the Oxford SA strains using a microarray that contained more than 3,600 PCR products to cover all potential open reading frames on the SA genome [114]. The core genome of each strain defined by the microarray correlated well with the MLST derived CC designations. There were no overt differences in the genomes defined by
microarray between invasive SA strains and carriage strains. MLST is able to differentiate two STs by a single point mutation, whereas these minor differences were not detected by the microarray. Microarray analysis confirmed that a significant proportion of each genome was made up of core genes that varied between different strains of SA. However, it revealed a large amount of variation in the mobile genetic elements: pathogenicity islands, resistance elements and bacteriophages.

In summary, increasingly detailed analyses of the Oxford SA isolates failed to identify any differences in the proportion of invasive SA isolates and nasal carriage isolates within each CC. However, they did identify bacterial factors that are important in the development of invasive disease: seven virulence genes were more likely to be found in invasive isolates than in nasal carriage isolates. Genomic analysis by microarray failed to identify any genetic elements that were more common in the invasive isolates than in the nasal carriage isolates.

5.1.2 Differences in invasive SA isolates and nasal carriage isolates of SA

Several other groups have looked for differences between invasive isolates of SA and nasal carriage isolates. Fowler et al compared the MLST and spa types of three groups of SA isolates: those from nasal carriage [n=118], those that caused uncomplicated SAB [n=56], and those that caused SAB associated with infective endocarditis or bone and joint infection [n=157] in North Carolina, USA. They found that strains of SA within CC5 and CC30 were more likely to cause severe invasive disease; although strains from all CCs were present in all three disease groups.

Melles et al compared of nasal carriage isolates and invasive disease isolates by amplified fragment length polymorphism [AFLP] and demonstrated that nasal carriage isolates were not randomly distributed amongst a sample of SA isolates from Rotterdam. However, the results of AFLP might be skewed by the presence or absence of mobile genetic elements. Van Belkum et al studied nasal carriage isolates [n=204] and SAB isolates [n=180] collected from people in Holland in the early 2000’s. The proportion of carriage strains that harboured the enterotoxin gene cluster mobile genetic element [64%] was significantly greater than the proportion of invasive strains that harboured the enterotoxin gene cluster [53%].

5.1.3 The global population structure of SA

The SA MLST database provides a rapidly growing representation of the global population SA; in late December 2009 the MLST database contained 3013 isolates of more than 1400 different sequence types. However, the database does not contain a completely representative sample of the global SA strains. In particular isolates from the Pacific constitute a small fraction of the total.
There are no isolates in the MLST database from the South Pacific, with the exception of four SA isolates from New Zealand; three of these four strains are methicillin-resistant SA [MRSA]. Likewise, a high proportion of the strains entered into the database from Australia are MRSA [129/178, 72%]; and only 7/178 [4%] of the entries were from New South Wales, Australia’s most populated state. Furthermore, in recent years it was very unusual for non-novel strains from Australia to be entered into the database. In 2008 only five isolates from Australia were entered and all were novel. It is probable that only strains unable to be genotyped by other methods were studied by MLST, or that only novel strains were entered into the global database.

Recent studies have avoided the potential for bias present in the MLST database by studying systematically collected strains from diverse locations around the world. The genetic population structure of SA from Dutch nasal carriers was found to be surprisingly similar to that of American nasal carriers. 30/300 [10%] of the STs of nasal carriage strains collected in France [n=81], Moldova [n=85], Algeria [n=85] and Cambodia [n=49] were novel [previously unrecorded in the MLST database]. Six CCs [CC5, CC8, CC15, CC30, CC45 and CC121] had disseminated to all sites, although the proportions of SA isolates within each CC varied between sites. There was more similarity between the French and Algerian samples of SA compared with the Cambodian sample. The Moldovan sample was almost equally related to the French and Algerian samples. The authors conclude that their findings suggest that modern human migration, which is frequent between Algeria and France and infrequent between Cambodia and other countries, accounted for global admixture of SA populations. A study of SA strains recovered from nasal carriers in Mali found that the majority belonged to globally disseminated CCs, but a divergent ST [ST152] was present in almost a quarter of the nasal carriers. The authors noted that ST152 was phylogenetically distinct from the main SA population. The global dissemination of SA was also evident from analysis of a sample of nasal carriage strains [n=147] obtained from children in China. This sample contained the same CCs found in other countries but the proportion of strains in each CC in China was more similar to the proportion of strains in each CC in Cambodia than in sample from Europe.

Regional variation in the proportion of strains within the most common CCs of SA is shown in Figure 5.1. The proportion of strains within each CC in European countries is similar. In contrast the USA sample contains a higher proportion of CC1 strains. The Chinese and Cambodian samples are similar; they each have a low proportion of CC30 strains and a high proportion of CC121 strains. There is a striking difference between the proportion of strains in each CC in Mali and the proportion of strains in each CC in neighbouring Algeria, where the proportion of strains within each CC is more similar to the European countries. This difference could be due to the collection of the SA sample in Algeria in a hospital close to the Mediterranean coast, a long distance from the Algeria-Mali border. Study of SA in New Zealand and the Pacific, will add greatly to the understanding of the population of SA globally; in particular whether or not neighbouring small Pacific Island nations display regional similarity in their populations of SA.
5.1.4 Population genetic study of \textit{S. aureus}

The nucleotide sequence data utilised by MLST is well suited to compare SA populations in different environments such as the SA populations which cause nasal colonisation or invasive disease; the SA populations that cause infections in different countries around the world; and, in the current study, the SA populations that cause infection in different ethnic groups within the Pacific region. Population genetic analysis of the nucleotide sequence data used to derive each SA strain’s ST aims to elucidate the evolutionary processes that shape each SA population: mutation, drift, recombination, selection, and migration.

5.1.5 Mutation, random genetic drift and phylogenies of SA

Mutations arise due to errors in DNA replication. While selection eliminates deleterious mutations and causes beneficial mutations to increase in frequency, random genetic drift determines the fate of neutral mutations. Many of the mutations that become fixed in a population by genetic drift are synonymous [mutations which do not encode a change in the amino acid sequence of a protein]. Mutations that are not fixed in the total population and are differentially inherited in lineages of SA define the phylogeny or ancestral tree of SA.

5.1.5.1 \textit{eBURST: sequence types and clonal complexes}
eBURST uses a simple and logical approach to understand the phylogeny of MLST data. A founder [or ancestral] ST increases in frequency, perhaps due to a fitness advantage, and from this type new variants arise, by either mutation or recombination. Should a new variant arise by mutation at one of the seven housekeeping loci, then this descendent can be distinguished by MLST and is referred to as a single locus variant [SLV]. As additional time passes mutations continue to arise such that variants arise from the SLV generating double locus variants [DLVs] that differ from the original ancestor at two of seven housekeeping loci. It is unlikely that a DLV will form spontaneously due to the slow rate of change in housekeeping genes. A CC is made up of an ancestor and its SLVs and DLVs [Figure 5.2]. One of the strengths of eBURST, which reconstructs the phylogeny from the allele designation, is that it does not matter how new variants arise. eBURST is good at grouping closely related strains [≥five of seven loci identical], which have diverged recently from a common ancestor. However, eBURST is unable to infer their evolutionary relationship for strains that differ by more than two loci [i.e., are identical at four loci]. In order to understand the evolutionary relationship between CCs, phylogenetic analyses that use the nucleotide sequence are required to provide a more detailed model of the ancestral relationships within a sample.

**Figure 5.2** The positions of the seven MLST loci within the SA genome; the arrow indicates the origin of replication. Genome B varies at the *arc* locus [yellow] and is a single locus variant [SLV] of A. C is a SLV of B and is a double locus variant [DLV] of A; A, B and C form a clonal complex. By eBURST analysis, D is unrelated to the other genomes, despite sharing common ancestry at 4 of 7 loci with C.

### 5.1.5.2 Phylogenetic tree reconstruction

There are a number of methods available to reconstruct phylogenetic trees from nucleotide sequence data. In broad terms these consist of methods that use a distance matrix to infer the tree [e.g. neighbour joining, NJ] or those that use the nucleotides [character state] to infer the tree. There are three commonly used methods to find the “best” tree using character state methods: maximum parsimony, maximum likelihood and Bayesian inference. Phylogenetic tree
reconstruction usually requires specification of a mathematical model [evolutionary distance] of nucleotide substitution.

5.1.5.3 Genetic distance and evolutionary distance: jMODELTEST

For homogenous SA populations the genetic and evolutionary distances are likely to be similar, but most phylogenetic analyses require specification of the evolutionary model of nucleotide substitution. Phylogenetic inference varies depending on the model of nucleotide substitution that is chosen. The pairwise genetic distance [PWD] is a measure of the nucleotide differences between two nucleotide sequences divided by the total length of the sequences. The genetic distance underestimates the evolutionary distance when divergence in the population is high. The genetic distance does not allow for the possibility that a nucleotide might have changed more than once during the evolutionary history. Evolutionary models take this, and other parameters, into account.

The simplest evolutionary model is the Jukes-Cantor model, which is not likely to be appropriate for most populations. This model specifies that substitution of one nucleotide for another occurs at the same rate and with equal frequency. More complex models allow for variation in the rate of transitions [purine to purine; pyrimidine to pyrimidine] and transversions [purine to pyrimidine and vice versa]. For example, the Hasegawa, Kishino and Yano model [HKY85] allows for different nucleotide frequencies and also for the differences in the rates of transitions and transversions.

jMODELTEST determines the most suitable nucleotide substitution model, from the large number of nucleotide substitution models that have been used for analysis nucleotide data. jMODELTEST calculates a number of likelihood values for a tree created using each of the models. When complicated evolutionary models are used to analyse genetic data computational time increases substantially; furthermore, complex models increase the error in each analysis as more parameters need to be estimated.

5.1.5.4 Phylogeny reconstruction using maximum likelihood: TREEUZZLE

Phylogenetic reconstruction is a crucial component of population genetics, following specification of an appropriate evolutionary model of nucleotide substitution. TREEUZZLE uses quartet puzzling to find the maximum likelihood [ML] tree. Initially the model parameters, such as nucleotide substitution model and total branch length, are optimised. Then the likelihoods of all of the possible quartets, the smallest group of taxa for which more than one rooted topology exists, are estimated. TREEUZZLE identifies quartets that are partially resolved [two topologies give similar likelihood scores] and unresolved [no topology is likely]. A large number of temporary trees are constructed by adding the taxa in a random order; taxa are added to a branch according to the quartet puzzling likelihood. A final majority rule consensus tree is derived from the temporary trees; the branch support is the proportion of times that branch occurred in the temporary trees.
5.1.5.5 Phylogeny reconstruction using a Bayesian method: CLONALFRAME

CLONALFRAME is a Bayesian method of tree reconstruction robust in the presence of modest rates of recombination and well suited to the analysis of MLST data. Bayesian methods generate a large set of possible trees and the posterior probability distribution is used to find the best tree for the data and the model used. CLONALFRAME identifies recombination events in the nucleotide alignment and estimates the tree for the portion of the nucleotide sequence that has not undergone recombination [the “clonal frame”]. Obviously, CLONALFRAME is not well suited for analysis of populations that freely recombine.

5.1.5.6 Estimating the demographic history of SA: GENIE

A well constructed phylogeny contains information beyond the evolutionary relationship of the taxa. Under coalescent theory the demographic history can be estimated. Coalescent theory refers to mathematical models that trace a phylogeny back in time from current taxa [tips of a phylogenetic tree] to the most recent common ancestor [hypothetical nodes of a phylogenetic tree]. The basic coalescent theory uses a model of neutral selection [random point mutation] to derive information about the ancestry of the sample. In the absence of selection pressure [that might favour some individuals over others] the time [backwards from the present] that it takes for two individuals to coalesce into a common ancestor is dependent on size of the population. A small population gains diversity more rapidly than a larger population. For example, if the size of a population has declined over time, then recent coalescent events will be rapid and further back in time coalescent events will be slower as the population increases in size. The reverse is true for a population that has increased in size. It is only possible to calculate the true ancestral population size for populations that are of constant size, which is not biologically plausible. Instead, the effective population size [Ne] is used, this measure reflects the rate at which genetic diversity is lost or gained.

The branch lengths of a phylogenetic tree provide a measure of the number of nucleotide substitutions that separate any pair of ancestral and daughter nodes. If the rate of substitutions is constant and if the molecular clock rate [the number of point mutations/unit of time] is known then the branches’ lengths can be converted to units of time. Thus, using a phylogenetic tree, it is possible to estimate changes in Ne over time; graphs of these measures are called “skyline plots”. The “classic skyline plot” allows Ne to change anytime there is a coalescent event in the phylogeny.

GENIE uses a likelihood process to estimate Ne at various times in the history of a phylogenetic tree and to reconstruct a skyline plot. It requires the input of a phylogenetic tree that has been formed under the assumption of the molecular clock. This assumption ensures that the lengths of the branches are proportional to time; the branch lengths between any two taxa and their common ancestor will be equal. GENIE is able to determine the likelihood of six demographic models; the most likely model is used to construct the demographic history.
5.1.6 Recombination

Recombination is a fundamental process in evolutionary biology and in bacterial populations occurs as a result of three para-sexual processes [transformation, transduction, conjugation]. The presence of recombination can alter the evolutionary relationships that have been inferred using many phylogenetic analyses; most of which are best suited to a model of evolution that does not include recombination. Detection and estimation of recombination is reliant upon detection of linkage disequilibrium [or allelic association], the lack of non-random sharing of variant nucleotides. The presence of allelic association between otherwise unrelated SA strains provides evidence of a recombination event during their evolutionary history.

5.1.6.1 Estimating the rate of recombination: LDHAT

In this thesis LDHAT was used in preference over the other methods and software packages available to estimate the rate of recombination within a population \(^{133-136}\). LDHAT was chosen because it employs several tests to detect the presence of recombination and it also quantifies \(\rho\), a parameter directly related to the degree of recombination present. Under the assumption of no recombination, polymorphic sites are exchangeable; they are independent of the order that they appear. If recombination is present, this independence is lost. LDHAT estimates the recombination rate of a population: \(\rho = 2N_e r\) where \(N_e\) is the effective population size and \(r\) is the rate at which nucleotides in the sequence differ by recombination. The mutation rate is \(\theta = 2N_e \mu\) where \(\mu\) is the mutation rate per nucleotide. The ratio of recombination to mutation is easily derived: \(\rho + \theta\). The LDHAT model assumes that the mutation rate is uniform across polymorphic sites and the first step estimates \(\theta\) using Watterson's estimate, and identifies every pair of polymorphic sites \(^{134}\). The next step involves estimating the likelihood of a range of recombination rates for each pair of polymorphic sites using the initial estimate of \(\theta\). A likelihood value is generated for each of the large number of possible independent genealogies for each pair of polymorphic sites across a range of recombination rates \(^{137}\). The final step provides the estimate of the population recombination rate by combination of the likelihood values of the comparisons of all of the polymorphic sites: the composite likelihood. The best estimate of \(\rho\) is the value that gives the highest composite likelihood.

LDHAT was also used to test the hypothesis that there was no recombination present \(^{134}\). LDHAT uses a likelihood permutation test [LPT]; whereby polymorphic sites are shuffled 1000 times and the composite likelihood is calculated from the random iterations. If no recombination has occurred then the composite likelihood value will not be dependant on the order in which sites occur and shuffling will lead to likelihood values similar to the true value for the data. Conversely, if recombination has occurred linked sites will be present and independence in the order of sites is lost; the likelihood values will be less than the true value for the data.

As no single test for recombination is reliable for all datasets, the PHI [pairwise homoplasy index] test was also used to detect the presence of recombination \(^{138}\). This test calculates the PHI
statistic, a measure of the similarity between neighbouring polymorphic sites; and was implemented in SPLITSTREE 4.2. Values of PHI are calculated for the alignment in blocks of 100 nucleotides. If a recombination event has occurred in the evolutionary history of the block of nucleotides, polymorphic sites that are close together will be more closely correlated than distant sites. If no recombination has occurred, then correlation between polymorphic sites will not be different for all of the sites. Statistical significance is determined by calculating a normal probability distribution for PHI.

5.1.6.2 Identifying sites of recombination: GARD

Recombination breakpoints were identified using the Genetic Algorithm for Recombination Detection [GARD]. GARD uses a phylogenetic approach to detect breakpoints. If recombination has not occurred within a population then all of the nucleotide sequences in the alignment are adequately represented by a single tree. GARD splits the nucleotide alignment into two partitions at each polymorphic site and then forms a neighbour joining [NJ] tree for each block. The likelihood measure of each new NJ tree topology and branch lengths are then compared to the NJ tree that represents the entire alignment. If the likelihood of the new NJ trees are greater than the original tree, then recombination has occurred at that polymorphic site in some of the sequences in the alignment. The Kishino Hasegawa topological incoherence test [KH test] is used to compare the topology of the NJ trees formed from the alignment upstream and downstream of each breakpoint. The KH test indicated a difference in tree topology if the p value was <0.025 [to correct for dual comparisons] between the two new trees and the original tree.

5.1.7 Selection pressure

Selection pressure operates on mutations that affect the phenotype of the organism. It is likely that populations of SA are so well adapted to their environments that most non-synonymous mutations [mutations which encode a change in the amino acid sequence of a protein] are detrimental and removed from the population by negative selection.

5.1.7.1 Detecting sites under selection pressure: SLAC

A synonymous change in a nucleotide sequence is one that encodes a protein with no alteration in the amino acid sequence. A non-synonymous change in a nucleotide sequence does encode a protein with an alteration in the amino acid sequence. In the absence of selection pressure, the rate of non-synonymous [dN] change is similar to the rate of synonymous change [dS] and the dN/dS ratio will be close to 1. Deviation in this ratio for a given nucleotide sequence alignment indicates the presence of positive [diversifying] selection [the ratio will be > 1] or negative [purifying] selection [the ratio will be <1].
A number of methods for detecting selection exist, in this thesis the Single-Likelihood Ancestor Counting [SLAC] method was used\textsuperscript{143,144}. SLAC reconstructs the hypothetical ancestral sequences for the taxa of a phylogenetic tree. Each hypothetical sequence is estimated using a maximum likelihood approach, and the translated amino acid sequence is derived to determine sites where a non-synonymous or synonymous change has occurred in the evolutionary history. Then the number of non-synonymous and synonymous substitutions at each site is calculated from the sequence alignment. \(dN\) is calculated as the number of non-synonymous changes that have occurred at a non-synonymous site and \(dS\) is calculated in the same way. In situations where multiple substitutions are possible along a branch of the tree, the number of synonymous and non-synonymous sites is taken from the average of the possible substitutions that does not include stop codons. A test of whether \(dN\) is significantly different to \(dS\) is performed by calculating \(p\) values from the assumption that the number of substitutions follows the binomial distribution. Kosakovsky et al found that using a \(P\) value cutoff of 0.1 gave low rates of type I error\textsuperscript{143}.

The rate at which mutations become fixed in a population is dependent on the effective population size and changes in the diversity of a population over time enable estimation of changes in population size over time. A small population gains and loses diversity rapidly; a large population gains or loses diversity slowly.

5.1.8 Migration

Migration allows the transfer of favourable genetic changes from one population to another. Under frequent migration the SA population in one country would be similar to another; infrequent migration causes geographically isolated populations to become dissimilar over time. Migration also allows SA strains to enter a new population and become successful in a new niche. The Western Samoan phage pattern methicillin-resistant SA strain [WSPP MRSA] has dispersed globally\textsuperscript{54,145}. As expected the dispersal of WSPP MRSA has been facilitated by human migration, WSPP MRSA in Alaska was associated with migration of Samoan families to Alaska\textsuperscript{38}. The comparison of the genetic structure of SA populations between countries [Figure 5.1] and between ethnic groups primarily concerns the migration and dispersal of SA strains.

5.1.8.1 Comparison of SA populations collected in different environments: MOTHUR

Three tests to compare SA sampled from different environments were utilised in MOTHUR\textsuperscript{146}. Firstly, a visual perspective was obtained by creating a Venn diagram. Secondly the Parsimony method was used\textsuperscript{147}. This test uses a Neighbour Joining tree of a population of SA to calculate the minimum number of migration events [parsimony score] between two sub-populations within the tree. The ST labels of the tree are then shuffled 1000 times and the parsimony score of each tree is then calculated to determine the proportion of random trees that have a parsimony score lower than or equal to the true tree. This random shuffle allows the estimation of a permutation
test p value. This test is extremely sensitive and even subtle variations in the STs between two populations give significant results. Thus, the Parsimony test is most useful if no significant differences are found between two populations, thereby avoiding the need for further testing.

The parsimony test detects differences in the evolutionary history between two groups while the LibShuff method compares the genetic similarity between two groups \(^{148}\). LibShuff calculates the coverage versus genetic distance from a distance matrix for each subpopulation. The coverage is calculated as: \(1 - \frac{n_i}{N}\) where \(n_i\) is the number of OTUs that have been sampled within distance \(i\) and \(N\) is the total number of strains in the sample. For example, if a sample of 10 strains contains 4 OTU\(_{unique}\)s the coverage at a distance of zero would be 60%. As the distance increases the coverage also increases until it reaches 100%. The difference in coverage at each distance is calculated between each group, and a statistic \(\Delta C\) is calculated to summarise the difference between the two curves. Finally, \(\Delta C\) is calculated for 10,000 random shuffles of the two groups to determine how often random permutations gave a value of \(\Delta C\) greater than the true value. If there are significant differences between groups then the random shuffle should reduce this difference most of the time and give values of \(\Delta C\) lower than the true value.

5.1.8.2 Non-phylogenetic visualisation of the genetic similarity between groups of SA: multi-dimensional scaling

Non-metric multi-dimensional scaling [MDS] displays genotypes of SA from different populations on a 2 [or 3] dimensional map \(^{149}\). The mathematical process behind MDS is beyond the scope of this thesis, but the interpretation is simple: the complex relationship between different genotypes is presented on axes [the axes have no specific meaning or value] preserving the relationship between members of the population. Genotypes that are close together on the MDS plot are closely related and those that are further apart are less closely related. MDS uses a PWD matrix to construct a map of the data derived for two dimensions and regression is performed on the distances between points on the axes. The ‘goodness of fit’ [stress value] is then calculated and the points are moved [maintaining reference to the other points] to reduce the stress required to represent the data in two dimensions. The optimisation procedure is repeated until no further improvement in the stress value is obtained. A stress score of <0.1 suggests an accurate portrayal of the data, unlikely to be improved by solutions that use more than two dimensions.

5.1.8.3 Comparison between populations: hypothesis testing using ANOSIM and PERMANOVA

ANOSIM and PERMANOVA perform statistical comparison of populations of SA displayed in 2-dimensions by MDS. ANOSIM [analysis of similarities] performs a non-parametric test which is analogous, in some respects, to the ANOVA [analysis of variance] test and was implemented in PRIMER 6 \(^{150}\). The test statistic \(R\) is a measure of the difference between two samples compared to differences within each sample. The \(R\) value between 0 and 1 is derived from the ranks of [dis]similarities in a matrix of genetic distance for each site and between each site, which is then scaled according to the population size. \(R = 0\) is found if the null hypothesis of no
difference is true; and \( R = 1 \) will be found if all genotypes within a site are more similar than any genotypes from different sites. A permutation test is performed, and the labels of the samples are shuffled and \( R \) is recalculated. For small data sets, it is possible to perform every possible permutation, for larger datasets 10,000 random permutations are performed. If the null hypothesis is true and \( R \) is close to 0, then changing the labels will not change \( R \). If the null hypothesis is false, and there is a significant difference between samples then the random permutation test will have the effect of reducing the difference between samples and \( R \) will usually be lower than the true value.

ANOSIM was used for one-way [univariate] tests, but for tests of more than one variable PERMANOVA was used.\(^{151}\) PERMANOVA partitions a genetic distance matrix into variables of interest [i.e. into groups] and determines the sum of squared distances within groups [\( SS_{\text{residual}} \)] and among groups [\( SS_a \)]. \( SS_{\text{residual}} \) is the distance between each member of the group, divided by the number of members. \( SS_a \) is the difference between the sum of squared distances for the total sample and \( SS_{\text{residual}} \).

A test of the null hypothesis is performed by calculating a pseudo-\( F \) statistic:

\[
\text{pseudo-}F = \frac{SS_a / [a-1]}{SS_{\text{residual}} / [N-a]}
\]

where \([a-1]\) is the total number of comparisons [degrees of freedom] associated with the variable and \([N-a]\) is the residual degrees of freedom.

Unlike the \( F \) statistic in traditional ANOVA, the pseudo-\( F \) statistic does not have a known distribution for the null hypothesis. Thus, like ANOSIM, the null hypothesis is tested by calculating pseudo-\( F \) for 10,000 permutations of the labels of the population members. If the null hypothesis is correct, and there is no difference among groups [i.e. no effect of the variable of interest], then the samples could have been obtained from any group and the pseudo-\( F \) value for each permutation will be similar to the true value. Conversely, if a difference does exist among groups then the value of pseudo-\( F \) will decrease as the labels are shuffled.

5.1.8.4 Determining populations of SA based on ancestral source: STRUCTURE

STRUCTURE was used to examine whether additional information about the ancestral source of different populations provides additional insight beyond that obtained from phylogenetic and non-model based analyses.\(^{152,153}\) For example, to determine if geographical clustering of SA populations was present, but was not evident by phylogenetic analysis. STRUCTURE has been used to identify the animal source of enteritis caused by campylobacter;\(^ {154}\) and has also been used to trace human migration over several hundred years, based on a modern sample of Heliocobacter pylori collected from diverse locations.\(^ {155}\)

STRUCTURE assigns individuals to populations based on their ancestry and has even been able to determine the ancestry of bacteria under free recombination.\(^ {155}\) The primary assumption used
by STRUCTURE is that alleles within populations are at complete linkage equilibrium; STRUCTURE identifies linkage disequilibrium and accounts for this by introducing population groups that are not in disequilibrium. A Bayesian approach is used to determine the number of ancestral populations and the frequencies of alleles in each population group. Each individual genotype is assigned to its ancestral population based on the proportion of linked alleles in each ancestral group.

5.1.9 Estimating the adequacy of the sample

The ability to draw meaningful conclusions from a population genetic study is reliant on the quality of the sample collected for the study as much as it is reliant on the careful application of robust analytical techniques.

5.1.9.1 Rarefaction and the reward of additional sampling

In this thesis, the adequacy of the sample was estimated in two ways. Firstly, rarefaction curves were drawn to compare the number of samples required to form the total number of operational taxonomic units [OTU] in the population.

An OTU contains all of the strains that differ by no more than a pre-specified distance. For example if each OTU was defined at a distance of 0.1, then each OTU will include only strains that have a PWD of <0.1; and strains that differ by a greater distance form other OTUs.

Rarefaction is performed by drawing sequential samples at random from the population, without replacement, to determine the number of OTUs generated for the number of samples drawn. This procedure is repeated 1000 times and the average estimate is used to generate a graph of OTU versus number of samples drawn. The rarefaction curve provides a visual estimate of the saturation of the sample for a given OTU, thus indicating the richness and diversity of the sample.

Secondly, the adequacy of the sample was estimated form the number of additional strains that would have been identified by additional sampling using Efron’s calculation:

\[
S_{\text{additional}} = \sum_{i=1}^{f_{\text{max}}} (-1)^{i+1} (M/N)^i f_i
\]

Where M is the number of additional strains collected and N is the total number of strains in the current sample. f is the number of strains in each OTU and the sum ends at \( f_{\text{max}} \) the largest OTU.

5.1.9.2 Comparison of diversity between populations
Simpson’s index of diversity was calculated to compare the diversity of SA sampled from different locations. The calculation takes the number of OTUs, the number of members of each OTU and the total sample size into account.

\[
\text{Diversity}_{\text{Simpson}} = \frac{\sum_{i=1}^{S_{\text{obs}}} n_i(n_i-1)}{N(N-1)}
\]

Where \( S_{\text{obs}} \) is the number of OTUs and \( i \) is the number of individuals in each OTU and \( N \) is the number of individuals in the population. Populations with greater diversity have lower values of Simpson’s index.
5.2 Aims

Three samples of SA were collected for the current study. The results of this Chapter are divided into three parts.

The first part of the results compares isolates of SA collected from cases of SAB [described in Chapter 1] and healthy nasal carriers [described in Chapter 4] in Auckland. This part aims to describe the genetic diversity and STs of the isolates of SA collected in Auckland during the study; to estimate the rates of recombination and point mutation of SA compared to other pathogens that cause nasopharyngeal colonisation and invasive disease; to examine the phylogeny of SA in Auckland; to describe the evolution of novel SA STs in Auckland, and to identify specific examples of mutation, recombination and migration that shape the evolution of SA in Auckland; to examine whether sub-populations of SA, defined by ancestral source or phylogenetic lineage, cause infections in different ethnic groups in Auckland; and to examine whether analysis of putative virulence gene sequences, in addition to concatenated MLST housekeeping genes, aids resolution of the genetic structure of SA populations.

The second part of the results compares samples of SA obtained from Fiji, Samoa and Tonga, with the sample of SA obtained in Auckland. This part aims to examine whether the genetic structure of populations of SA in neighbouring Pacific Island countries are similar; and whether some lineages of SA in the Pacific that have a predisposition to cause infection in people of different ethnicities.

The third part of the results compares the genetic structure of the population of SA in the Pacific with the SA populations previously described elsewhere in the world: Oxford, Mali and China. This part aims to examine the similarity of SA populations obtained from people living in extremely diverse locations and to find evidence of a recent population expansion of SA that mirrors the recent expansion in the human population.
5.3 Materials and Methods

5.3.1 Identification and testing of SA isolated from nasal swabs

The strategy of identification and sampling of SA from nasal swab specimens is summarised in Figure 5.3. Identification and antimicrobial susceptibility testing of SA isolated from blood cultures was performed by the staff of microbiology laboratories at Auckland City Hospital and Middlemore Hospital.

Nasal swab specimens were swabbed directly onto mannitol salt agar [Fort Richard, Auckland]. Colonies of SA were presumptively identified as SA if they utilised the mannitol in the agar and produced a carbohydrate acid that turned the surrounding agar from pink to yellow. Identification of SA was confirmed by suspending a few colonies of the isolate in the Staphyloslide™ Latex test kit liquid [BD, Mt Wellington, Auckland, NZ] containing latex particles coated with human fibrinogen and IgG; leading to agglutination of the latex particles by protein A and/or clumping factor present on the surface of the SA cells. This test has a sensitivity of >94.3% and a specificity of 98% \(^{156}\). The identity of SA was unequivocally confirmed by testing for the expression of coagulase and DNase.

To perform the coagulase test, one to two colonies of the isolate were placed in a tube containing rabbit serum and then incubated at 37°C for at least 16 hours and examined for clot formation. Testing for expression of DNase was performed by plating 5uL of a 1:5 dilution of 0.5 McFarland standard suspension onto DNase test agar, incubated at 37°C for 16 hours. The enzyme reaction was catalysed by the addition of 0.5 M HCl, washed from the plate after 30 seconds. A positive test was indicated by a zone of clearance of the nucleic acid in the agar surrounding the bacterial colony. The reference strain of S. epidermidis ATCC12228 was used as a negative control.
The combination of coagulase and DNase test is sensitive and specific for the identification of SA. On the rare occasion when conflicting results were obtained, the API Staph ID32 [bioMérieux SA, Auckland, NZ] was used to confirm or refute the identity of SA. The ID32 consists of a strip containing reagents for 26 colorimetric biochemical tests. A drop of bacterial suspension of 0.5 McFarland standard in sterile saline was inoculated to each well of the strip. APILAB software was used to interpret the results of the 26 tests and the identity of SA was confirmed if supported by a probability of \( \geq 98\% \).

Antimicrobial susceptibility was determined by agar breakpoint dilution testing for oxacillin, cotrimoxazole, doxycycline, gentamicin, amikacin and vancomycin using a 1:5 dilution of a 0.5 McFarland standard suspension of the test organism. 5uL of sample was placed onto agar containing a known concentration of test antibiotic in a grid with other samples. Each plate was incubated at 37\(^\circ\)C for at least 16 hours and then examined for the presence of bacterial growth. The results were interpreted in accordance with criteria defined by the Clinical and Laboratory Standards Institute methods and interpretive standards. An in-house reference strain of MRSA [oxacillin MIC=8 mg ml\(^{-1}\)] was used as a control. Penicillin susceptibility was tested by disc diffusion.
5.3.2 Selection of SA colonies for genotypic testing

To avoid potential bias influenced by the appearance of the colonies a random sampling process was used to select SA colonies from the agar plates that had been inoculated either with broth from positive blood cultures or with nasal swabs. A 3x3 grid was placed behind the agar plate and a random number generator was used to select one of the nine grids. The single SA colony closest to the centre of the selected grid was sampled using a sterile toothpick and the bacteria were placed into sterile storage broth [35% v/v glycerol, 50mM MgSO$_4$, 125uM TRIS pH8] and stored at -80°C. Two colonies were sampled in this manner from each plate.

5.3.3 Genomic DNA preparation

Frozen bacterial stocks were resuspended and incubated overnight in 7 ml of LB broth at 37°C. The following morning 1.5 ml of culture was centrifuged at 10000 rpm for two minutes and the supernatant carefully removed by sterile pipette. The pellet was then resuspended in 180uL lysis buffer containing 20 mg ml$^{-1}$ lysozyme and 25 ug ml$^{-1}$ lysostaphin. After 30 minutes of incubation at 37 degrees, 25 uL of proteinase K was added and incubated at 56°C for a further 30 minutes. Genomic DNA was then extracted using the DNeasy kit [Qiagen, Hilden, Germany] according to the manufacturer's instructions and stored at -20°C, for use in all PCR reactions.

5.3.4 Polymerase chain reaction [PCR] and PCR primer design

Multi locus sequence typing [MLST] of SA compares the nucleotide sequences from segments of seven housekeeping genes: arcC, aroE, glpF, gmk, pta, tpi and yqi. The oligonucleotide PCR primer sequences and conditions for their use were downloaded from http://saureus.mlst.net/ $^{57}$. For initial testing the primer pairs [Sigma Aldrich, NSW, Australia] for arcC and aroE were tested against a range of SA isolates under a range of PCR conditions. The primer pair for the aroE gene did not produce a product when trialled across a range of annealing temperatures, using a range of magnesium concentrations against a number of SA DNA templates. Following this result, seven new PCR primer pairs were designed.

The new primers were designed to amplify the nucleotide sequence amplified by the conventional primers against a range of staphylococcal species. Each MLST gene sequence from SA strains Mu50, MW2, N315, NCTC8325, RF122, COL, MRSA 252, MSSA 476, and Newman, together with S. epidermidis ATCC12228, S. haemolyticus JCSC1435 and S. saprophyticus ATCC15305 [from http://cmr.jcvi.org/cgi-bin/CMR/CmrHomePage.cgi] were aligned using Geneious 3.5 $^{158}$. New PCR primer pairs were designed to amplify a larger region that contained the consensus sequence plus adjacent flanking sequences.

PCR was performed using genomic DNA extracted from each of the two SA colonies recovered for each specimen. PCR was performed in 96 well plates with the total reaction in 25uL [1.5 mM
MgCl₂, 0.2 mM dNTP [Invitrogen, Auckland, New Zealand], 0.5 uM forward and reverse primers [Sigma Aldrich, NSW, Australia], 2.5 units Taq DNA polymerase [Invitrogen] in PCR buffer [Invitrogen]. Each reaction contained 2.5uL of DNA template. Amplification was continued for 35 cycles of displacement at 94 degrees for 1 minute, annealing for 30 seconds and extension at 72°C for 1 minute. The PCR products were confirmed by electrophoresis of 5uL of amplicon in 0.5% Ultrapure agarose [Invitrogen] stained using 1:10,000 SYBR safe DNA gel stain [Invitrogen]. The products were visualised and photographed under ultraviolet light at 302nm.

The optimal annealing temperature was selected as the one that produced the clearest single band of the predicted size when viewed on agarose gel. Each of the PCR primer pairs were then tested using DNA extracted from seven of staphylococcal reference strains: Oxford SA ATCC9144, SA ATCC25923, MRSA ST92, S. capitis ATCC27840, S. lugdunensis NHST87/1001, S. saprophyticus ATCC15305, and S. epidermidis ATCC12228 [Figure 5.4]. For the arcC gene, the conventional primers performed better than the new primers and were used in the study [Table 5.1].

Each PCR product was purified by adding 5uL of purification solution [4.7ul H₂O, 0.2ul Calf intestinal phosphatase, and 0.1ul Exonuclease I to each well followed by incubation at 37°C for 30mins and 85°C for 15 mins. Each 96 well plate was packaged in accordance with IATA protocols and couriered by air to Macrogen Inc [Seoul, Korea] for sequencing. At Macrogen sequencing was performed using microcapilliary electrophoresis of the purified PCR products using BigDye™ terminator cycling.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence 5’ to 3’</th>
<th>Annealing temperature °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ox arc fwd</td>
<td>TTG ATT CAC CAG CGC GTA TTG TC</td>
<td>55</td>
</tr>
<tr>
<td>ox arc rev</td>
<td>AGG TAT CTG CTT CAA TCA GCG</td>
<td>55</td>
</tr>
<tr>
<td>aro fwd</td>
<td>CCY ATT TCW CAT TCM TTR TCG C</td>
<td>54</td>
</tr>
<tr>
<td>aro rev</td>
<td>CAT ACC MGC WGG TGT WGT RTT</td>
<td>54</td>
</tr>
<tr>
<td>glp fwd</td>
<td>YTW TTY GGK GGT GGC GTT TGT GC</td>
<td>60</td>
</tr>
<tr>
<td>glp rev</td>
<td>SHC CAR TTW GAW YYM CCT TTV CC</td>
<td>60</td>
</tr>
<tr>
<td>gmk fwd</td>
<td>CCW TCW GGH GTW GGD AAR GG</td>
<td>50</td>
</tr>
<tr>
<td>gmk rev</td>
<td>CGY TCT CTY TTY AAR TGY TCA GC</td>
<td>50</td>
</tr>
<tr>
<td>pta fwd</td>
<td>GTN AAA ATM GTW TTA CCW GAR GG</td>
<td>54</td>
</tr>
<tr>
<td>pta rev</td>
<td>GAM CCT TTW GTW GAR AAG CTT AA</td>
<td>54</td>
</tr>
<tr>
<td>tpi fwd</td>
<td>GWC AYT CHG ARC GTC GTG AM</td>
<td>52</td>
</tr>
<tr>
<td>tpi rev</td>
<td>TTT GCA CCT TCT AAC AAT TGT AC</td>
<td>52</td>
</tr>
<tr>
<td>yqi fwd</td>
<td>MAG CAT AYM GRA CAC CTA TWG GC</td>
<td>58</td>
</tr>
<tr>
<td>yqi rev</td>
<td>CGT TGW GGW AYC KMT ACW GGW AC</td>
<td>58</td>
</tr>
<tr>
<td>16sRNA fwd</td>
<td>AGA GTT TGA TCM TGG CTC AG</td>
<td>52</td>
</tr>
<tr>
<td>16s RNA rev</td>
<td>AAG GAG GTG WTC CAR CC</td>
<td>52</td>
</tr>
<tr>
<td>ssl7 fwd</td>
<td>CAT TAG GCT TAT TAA CTA CTG GTG</td>
<td>56</td>
</tr>
<tr>
<td>ssl7 rev</td>
<td>CCG TTA TAR TTT CTA ACC TTA CCA C</td>
<td>56</td>
</tr>
<tr>
<td>coa fwd</td>
<td>GAY GMW TCT TGG AAA AMD ARA ACY GTC</td>
<td>54</td>
</tr>
<tr>
<td>coa rev</td>
<td>CCA AAT GTT CCA TCG TTG TAT TCA CGG</td>
<td>54</td>
</tr>
</tbody>
</table>

Table 5.1 The PCR primers and conditions used for SA MLST.
5.3.5 Sequence manipulation

Sequences were obtained for both DNA strands and they were assembled and aligned using Geneious 3.5. The portions of the sequences flanking the conventional MLST sequences were trimmed and the conventional MLST sequences were compared with the sequences in the MLST database [http://saureus.mlst.net/] to obtain the allele number for each locus. On the rare occasion when the DNA sequence for the paired strains did not match, the PCR and sequencing was repeated on DNA extracted from the original frozen bacterial stocks. On the more common occasion that the DNA sequences for the paired strains were identical but did not match a known allele in the MLST database, the gene sequence was confirmed by repeat PCR and sequencing of that gene for both of the paired isolates. The final sets of MLST sequences were concatenated into a total sequence of 3198 nucleotides for the majority of the analyses.

5.3.6 Novel SA strains

The designations of an ST number of 11100 to 11150 were used to identify STs that had not been identified previously in the SA MLST database [http://saureus.mlst.net]. Each novel ST was compared to the MLST database, to identify the most closely related known ST. Novel alleles arose by point mutation or by recombination when compared to the closest known ST [the
putative ancestor]. A recombination event was identified if a gene sequence differed by more than one nucleotide from the ancestor allele, or if a known allele was present in a different genetic background [Figure 5.5].

<table>
<thead>
<tr>
<th>Consensus</th>
<th>G A C T G C C T T C A C T G C T A G G C C T C T C C G G C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G A C T G C C T T C A C T G C T A G G C C T C T C C G G C</td>
</tr>
<tr>
<td>2</td>
<td>G A C T G C C T T C A C T G C T A G G C C T C T C C G G C</td>
</tr>
<tr>
<td>3</td>
<td>G A C T G C C T T C A C T G C T A G G C C T C T C C G G C</td>
</tr>
</tbody>
</table>

Figure 5.5 An example of mutation and recombination in comparison with a reference strain [1]. Strain 2 differs from strain 1 by a point mutation at position 15, a change considered to have arisen by mutation. Strain 3 differs from strain 1 at positions 8 and 24, changes most likely to have arisen following a recombination event.

5.3.7 Software and parameter settings

5.3.7.1 eBURST

Clonal complexes [CC] were defined using the eBURST algorithm implemented in goeBURST v1.2.1. goeBURST provides an additional step to the eBURST algorithm to resolve situations when the eBURST algorithm might fail. In addition it forms groups of clonal complexes [CC] and STs that are triple locus variants; that is, STs that are identical at four of seven alleles.

5.3.7.2 jMODELTEST 0.1.1

jMODELTEST 0.1.1 was used to identify the most likely evolutionary model for populations of SA. The concatenated MLST nucleotide sequences of all 244 SA strains obtained in Auckland were analysed; by hierarchical likelihood ratio testing [which compares the likelihood values of two different models] the most likely evolutionary model for the SA population was the general time reversible model [GTR] with a proportion of invariable sites [I] and with a gamma distribution for the variability in the rate of substitution at different nucleotides in the sequence [G].

A simpler model with a high likelihood [lowest AIC value] the HKY model, was used consistently in each analysis in the current study, in order to reduce computational time and because the GTR + I + G model was not available in all of the software packages used [Figure 5.6].
Figure 5.6 The likelihood [AIC] of 88 models used to construct a tree of the SA MLST data estimated using jMODELTEST 0.1.1. GTR=general time reversible, HKY=Hasegawa, Kishino and Yano, JC=Jukes-Cantor, I=proportion of invariable sites, G=gamma distribution of nucleotide substitution rates. Other model labels have been removed to aid the display of the data.

5.3.7.3 MOTHUR

The software package MOTHUR [http://www.mothur.org/wiki/Main_Page] was used to draw rarefaction curves. MOTHUR was also used to estimate the number of additional taxa, $S_{\text{additional}}$, if each sample was increased by 10%. The smallest pairwise distance that MOTHUR was able to differentiate and group into an OTU, was $<0.0049$. Rarefaction curves and $S_{\text{additional}}$ was estimated using OTU$_{\text{unique}}$ [each OTU contained identical strains] and OTU$_{0.0049}$ [strains that varied by a distance $<0.0049$ were grouped into each OTU].

MOTHUR was also used to compare populations using the Parsimony test and LibShuff. In LibShuff a $P$ value of $<0.05$ [i.e. a random shuffle gave a value of $\Delta C >$ the true value less than 500/10,000 times] was used to infer a statistically significant difference between groups.

5.3.7.4 LDHAT

In this thesis, the software package LDHAT [http://www.stats.ox.ac.uk/~mcvean/LDhat/] was used to estimate the rate of recombination and mutation. The LDHAT likelihood permutation test was used to detect the presence of recombination within a population. The test was considered statistically significant if less than 50 of the 1000 [$P \leq 0.05$] random permutations gave a likelihood value greater than that calculated from the original data. In order to maximise the power of this test, only polymorphic sites with two alleles and a minor allele frequency of more than 10% were analysed.
5.3.7.5 PHI test

The PHI test was used as a complementary test to detect recombination and was implemented in SPLITSTREE 4.2 [http://www-ab.informatik.uni-tuebingen.de/software/splitstree4] \(^{139}\). P values $\leq 0.05$ were considered significant.

5.3.7.6 GARD

Recombination breakpoints were identified using the Genetic Algorithm for Recombination Detection [GARD] \(^{140, 141}\). To identify the presence of multiple breakpoints GARD searches a very large number of trees and can be implemented in the HYPHY software package \(^{162}\). However, due to computational requirements, it was performed on a computer cluster hosted at the University of California via a web interface [http://www.datamonkey.org/].

5.3.7.7 PRIMER 6

Visual comparison of populations of SA was performed by multi-dimensional scaling implemented in PRIMER 6.1.12 \(^{149}\). Two dimensional maps were used; and the algorithm was restarted 50 times to ensure that the lowest stress value had been obtained.

Comparison between populations of SA based on genetic distance that did not require mathematical modelling, ANOSIM and PERMANOVA were also implemented in PRIMER 6 and PERMANOVA+ 1.0.2 \(^{151}\). In ANOSIM, statistical significance was inferred from a permutation test that gave an R value $\geq$ the true R value less than 5% of the time $[p=0.05]$. In PERMANOVA statistical significance was inferred if $\leq$5% of the permutations gave a pseudo-F value $\geq$ the true value. All permutations were performed on the raw data from genetic distance matrix.

5.3.7.8 TREEPUZZLE

TREEPUZZLE 5.2 [http://www.tree-puzzle.de/] was used to reconstruct maximum likelihood [ML] trees \(^{127}\). Each tree was reconstructed using the HKY85+G model with the nucleotide frequencies and transition to transversions ratios estimated from the data. An NJ tree was used initially to derive the parameter values and 100,000 puzzling steps were used for each tree. The TREEPUZZLE output was displayed in tree form manipulated using FIGTREE 1.2.2 \(^{163}\).

5.3.7.9 CLONALFRAME

Bayesian phylogenetic trees robust in the presence of moderate degrees of recombination were reconstructed using CLONALFRAME 1.1 [http://www2.warwick.ac.uk/fac/sci/statistics/staff/research/didelot/clonalframe/] \(^{128}\). The parameters required for the model were initially estimated over a number of iterations of the algorithm, starting with a random value. The first 50,000 iterations of the algorithm were discarded as “burn-in”: the number of iterations taken for the parameters to reach their best estimate. Each analysis was run for a minimum of 500,000 iterations, in duplicate, using different
starting values. Duplicate analyses were compared for convergence of values of several parameters such as mutation rate \([\theta]\), recombination rate \([\rho]\) and the total branch length. When the ratios of these values for separate runs approached 1.0 and when the tree topology was comparable a majority rule consensus tree was formed from the duplicate trees. The majority rule tree required 50% support for each branch; that is, the branch was present in \(\geq 50\%\) of the trees in the posterior distribution of the two runs.

If convergence in the parameters between two duplicate runs was not obtained after 500,000 iterations, then the number of iterations was increased [typically to 1,000,000] and the number of iterations among samples was doubled [initially to 200] until convergence was obtained. The CLONALFRAME output was displayed in network form using CYTOSCAPE 2.6.3\(^{164}\).

5.3.7.10 STRUCTURE

STRUCTURE 2.2 [http://pritch.bsd.uchicago.edu/structure.html] grouped SA isolates on the basis of their ancestry to determine if geographical clustering of SA populations was present that was not evident by phylogenetic analysis\(^{152, 153}\). The final three repetitions of each analysis were grouped together using CLUMPP 1.1.2 and then displayed visually using DISTRUCT\(^{165, 166}\).

STRUCTURE creates a map of polymorphic sites from each alignment of SA sequences. \(\lambda\), a parameter describing the frequency of alleles in each population, was estimated from two interim repetitions of the analysis. The admixture model was used for each analysis [experimentation found very little difference when the alternative “linkage” model was used for SA populations], which allows an individual to inherit some of her genome from different ancestral populations. The degree of admixture \([\alpha]\) was inferred from the model and the same value of \(\alpha\) was used for each population; the initial value of \(\alpha\) was set at 1. The number of ancestral populations \([K]\) was inferred from the genetic data, 3 iterations were performed for values of \(K\) between 1 and 12, run for 100,000 repetitions after a burn-in of 25,000 repetitions. \(K\) was selected as the value where the likelihood of the model stopped improving rapidly. Once \(K\) had been selected, the data output was collected from three iterations of 500,000 repetitions after a 50,000 repetition burn-in. To assign admixed individuals to an ancestral group, a cut off of more than 67% was used. If an individual did not have more than 67% of its genome inherited from a single ancestral group then it was placed into a new group of individuals with mixed inheritance.

5.3.8 The global dataset of SA MLST sequences

There were two potential sources of MLST sequence data from other parts of the world, the MLST database [http://saureus.mlst.net/] and the published literature. The MLST database has limitations and was not used for comparison with the sample from Pacific nations.

The published literature was searched for studies that have collected an ecologically sound sample and have performed MLST on all of the SA isolates that were obtained. It was also
desirable that these studies did not obtain their samples from hospital staff or hospitalised patients. Three studies were identified that met these criteria. The first was conducted by Day et al and included strains used in the development of the MLST scheme for SA\textsuperscript{57, 110, 112}. This study obtained isolates that caused invasive disease in England in 1997 and 1998, as well as isolates obtained from healthy blood donors. The second study obtained isolates of SA from asymptomatic nasal carriers within 8 hours of admission to hospital in Mali for emergency surgery\textsuperscript{119}. The third study obtained 198 disease and carriage isolates from children in Chengdu, Sichuan province, China\textsuperscript{107}.

Two further studies were considered for inclusion. One compared SA strains obtained from nasal carriers in several countries, but it was not possible to determine the STs identified in the paper or the MLST database\textsuperscript{120}. The other study obtained SA from nasal carriers who were healthcare workers, and acquisition in the community could not be confirmed\textsuperscript{167}.

The nucleotide sequences of the STs in the three community based studies were obtained from the MLST database. Only unique STs were included, because it was not always possible to determine the proportion of isolates in each ST. The English study contained two STs that were not included in the comparison dataset\textsuperscript{112}. ST21 contained a stop codon in the \textit{glp} gene and the nucleotide sequence or allelic profile of ST167 was not available in the MLST database.

The final global sample contained 73 STs isolated in England ten years ago; 33 STs isolated from children in China; 20 STs from nasal carriers in Mali and the 87 STs isolated in Auckland, Fiji, Samoa and Tonga as described in this thesis.
5.4 Results

5.4.1 The population structure of *S. aureus* that caused bacteraemia and asymptomatic nasal colonisation in Auckland during 2007

This section describes the population structure of the isolates of SA associated with invasive disease or asymptomatic colonisation of human hosts in Auckland, during 2007. The analysis focuses on comparison of the isolates that infect Māori and Pacific people, with the isolates that infect people from other ethnic groups.

5.4.1.1 The strains of SA isolated from people with bacteraemia

In order to collect strains from a wide range of Auckland’s society the study was conducted at two of Auckland’s three large public hospitals, during 2007. Thus, the total dataset contained SA strains isolated from 150 adult cases with SA bacteraemia [SAB]: 96 isolates from cases at Auckland City Hospital and 54 isolates from cases at Middlemore Hospital. The features of the cases are described in Chapter 1; a summary is shown in Table 5.2. All of the cases were identified following the growth of SA from blood culture. Attempts were made to recruit all consecutive cases to the study and no exclusion criteria were applied. Six potential cases were identified, but were not included in the study.

5.4.1.2 The strains of SA isolated from people with asymptomatic nasal colonisation

In order to compare these invasive isolates with the strains that caused asymptomatic nasal colonisation in the Auckland community, a further 94 strains isolated from healthy people were included. 78 strains were isolated from people enrolled in a study of population nasal carriage of SA; these people were described in Chapter 4.

<table>
<thead>
<tr>
<th>sample site</th>
<th>specimen</th>
<th>mean age +/- s.d</th>
<th>female</th>
<th>Māori</th>
<th>Pacific</th>
<th>Other ethnicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auckland City Hospital [96]</td>
<td>blood culture</td>
<td>57 +/- 20 years</td>
<td>31 [32%]</td>
<td>16 [17%]</td>
<td>22 [23%]</td>
<td>58 [60%]</td>
</tr>
<tr>
<td>Middlemore Hospital [54]</td>
<td>blood culture</td>
<td>53 +/- 20 years</td>
<td>17 [31%]</td>
<td>16 [30%]</td>
<td>22 [40%]</td>
<td>16 [30%]</td>
</tr>
<tr>
<td>Auckland Community [78]</td>
<td>nasal swab</td>
<td>26 +/- 10 years</td>
<td>34 [44%]</td>
<td>6 [8%]</td>
<td>8 [10%]</td>
<td>64 [82%]</td>
</tr>
<tr>
<td>Household contacts [16]</td>
<td>nasal swab</td>
<td>37 +/- 17 years</td>
<td>11 [69%]</td>
<td>0</td>
<td>4 [25%]</td>
<td>12 [75%]</td>
</tr>
</tbody>
</table>

Table 5.2 The characteristics of 244 participants who provided an isolate of SA from blood culture or nasal swab culture.
A further 16 isolates were obtained from household contacts of the patients with SAB. These swabs were taken to assess the prevalence of SA nasal carriage in households of cases of SAB and to determine the genotypic relationship with the bacteraemia strains. It was expected that the carriage rate in “infected” households would be high and that all of the carriage strains would be identical to the bacteraemia strain. Participants for this study were difficult to recruit and this study was abandoned. However, nasal swabs were obtained from 33 consenting household members who were present in the hospital during one of the study visits. 17/33 [52%] of the nasal swabs cultured SA, but one isolate could not be recovered for genotyping from frozen stocks. Only 8/16 [50%] were the same ST as the isolate that caused SAB in their household member; none of the 33 people who had swabs performed had recent healthcare contact and none had symptoms or signs of disease caused by SA. For these reasons, these isolates were included in the sample of SA obtained from healthy population members.

5.4.1.3 Polymorphism within the Auckland sample of SA

The degree of polymorphism at the nucleotide and amino acid levels for each of the seven housekeeping genes and for the concatenated sequence of seven genes is shown in Figure 5.7. The concatenated sequences contained 313 [313/3198, 10%] polymorphic nucleotides and 55 [55/1066, 5%] polymorphic amino acids. The yqi gene contained the most diversity with 13% polymorphic nucleotides and 9% polymorphic amino acids. The glp gene contained a large number of synonymous polymorphic nucleotides: 12% of the nucleotides but only 1% of the amino acids were polymorphic. The aro gene contained the fewest polymorphic nucleotides [20/456, 5%]; yet the number of polymorphic amino acids [9/152, 6%] was similar to the other genes.

![Figure 5.7](image-url)
5.4.1.4 Identification of a genetic outgroup of SA that caused nasal colonisation

The mean pairwise distance [PWD] between SA strains was 28 nucleotides per 3198 [PWD 0.0088] nucleotides. However, four strains isolated from nasal swabs from healthy population members formed a distinct outgroup. Three of these strains were identical and the other differed by only 21 nucleotides; however, these four outlier isolates differed from the main population of SA by up to 324 nucleotides [PWD 0.1013]. Figure 5.8 shows that the genetic diversity [defined by PWD] of the strains of SA that cause SAB [blood] and nasal colonisation [nasal] are very similar, with the exception of the outlier strains, represented by the four bars to the right of Figure 5.8.

The four outlier strains were investigated further to determine whether they represented a different species, or resulted from contamination of the PCR reaction. All four strains were re-cultured from frozen stocks and repeat phenotypic testing indicated coagulase and DNase positivity. Additional testing using Staphylolide™ Latex test kit liquid [BD, Mt Wellington, Auckland, NZ] and the API Staph ID32 [bioMérieux SA, Auckland, NZ] confirmed the identification of SA. There are some limitations with identification of staphylococci using metabolic tests; and so the identity of these strains was also confirmed by determining the 16sRNA sequences of these isolates. DNA was extracted from the colonies that had been re-tested phenotypically and the PCR amplification and sequencing of the seven housekeeping genes was repeated, confirming the initial sequence results.

![Figure 5.8 The frequency distribution, expressed as a percentage of the total, of the pairwise genetic distance of the Auckland SA population. Two sub-populations are shown; those that caused SAB [blood] and those found in healthy population carriers [nasal]. The two populations are very similar, but the nasal population contained four strains that differed from the main population by approximately 300 nucleotides.](image-url)
5.4.1.5 The Auckland sample of SA is a comprehensive representation of the total SA population in Auckland

Duplicate genotypes are removed prior to performing many population genetic analyses. Thus, it is important to establish that the sample provides an adequate representation of the total population of SA. This is also critical for comparison between populations of SA, for example between different countries or between different ethnic groups.

The 240 Auckland samples [excluding the outlier nasal carriage strains] were clustered into operational taxonomic units [OTUs] based on their genetic distance using MOTHUR. An OTU consists of a number of strains grouped together based on their genetic similarity. For example, each OTU\textsubscript{unique} only contained unique [and identical] strains; in the sample there were 60 OTU\textsubscript{unique}s [excluding the four outlier strains]. A rarefaction curve was obtained from 1,000 iterations of samples drawn at random to determine the sample size required to form each OTU\textsubscript{unique} [Figure 5.9]. The rarefaction curve for unique isolates was not saturated despite the sample size of 240, suggesting that substantial additional sampling is required to find all of the unique strains in Auckland.

![Figure 5.9](image.png)

The rarefaction curve for the Auckland sample of SA [n= 240]. The graph shows the curve of unique strains and strains with a distance of < 0.0049 are clustered into operational taxonomic units [OTUs] versus the number in the sample.

The smallest genetic distance that MOTHUR is able to differentiate is a PWD of $\leq$0.0049, which equates to 16 nucleotide changes in the concatenated sequences of 3198 nucleotides. The Auckland samples clustered into 13 OTU\textsubscript{0.0049} [OTUs differentiated by a distance of $\leq$0.0049]. This level of richness could have been found within a sample size of approximately 175. This is
the point on the rarefaction curve [Figure 5.9] where the number OTU\textsubscript{0.0049}s stops increasing despite an increase in sample size, and becomes saturated. The graph shows that after collection of 50 samples 11 OTU\textsubscript{0.0049}s were identified; after 100 samples 12 OTU\textsubscript{0.0049}s were identified. The 13 OTU\textsubscript{0.0049}s were identified after 175 samples, and further increase in sample size did not increase the number of OTU\textsubscript{0.0049}s. A large number of additional samples would have been required to increase the number of OTU\textsubscript{0.0049}s further. Clustering the isolates into OTU\textsubscript{0.0049}s provided a conservative measure of the rarefaction of the sample. The mean PWD in the sample was 0.009; clustering the isolates into OTU\textsubscript{0.0049}s provided a rarefaction curve that became saturated with far fewer isolates.

5.4.1.6 The sample of SA in Auckland would not have improved with additional sampling

The adequacy of the sample was also estimated by calculating the number of additional OTUs that would have been identified by additional sampling [Section 5.1.9.1]. Efron's calculation estimated that an additional 3 unique strains would have been identified if the number of people sampled was increased by 10%; no additional OTU\textsubscript{0.0049}s would have been identified. Thus, additional sampling would have had minimal impact on the diversity of the sample as a whole. The Shen and Solow formulae, also implemented in MOTHUR, gave the same result\textsuperscript{146}.

5.4.1.7 The rates of recombination and point mutation of SA isolated from people living in Auckland.

Prior to performing further analysis of the SA sample, it is important to determine whether recombination has occurred within [or between] the housekeeping genes. Recombination makes the inference of evolutionary relationships problematic. The wide separation of the MLST loci on the SA genome makes it unlikely that a single recombination event would affect more than one locus\textsuperscript{109}. However, phylogenetic analysis of MLST sequence data showed that the evolutionary descent of E. coli was altered when recombination events were taken into account\textsuperscript{169}. Bruen et al demonstrated that tests vary in their ability to detect recombination\textsuperscript{138}. Therefore, two tests: the LDHAT likelihood permutation test and the PHI test were used to look for evidence of recombination [Section 5.1.6.1]\textsuperscript{134, 138}.

5.4.1.8 Recombination was detected in nucleotide sequences of housekeeping genes of SA

The PHI test and LDHAT likelihood permutation test confirmed the presence of recombination in the concatenated sequences of the total population of 244 SA strains. In order to ensure that the recombination signal was not enhanced or obscured by the four outlier strains, analysis was performed on nasal population of strains with and without the outlier strains. Recombination was present in all the blood [from cases of SAB] and nasal [from healthy nasal carriers] strains of SA [Table 5.3].
After the presence of recombination was confirmed, the rates of recombination \([\rho]\) and mutation \([\theta]\) of the Auckland SA population were estimated. LDHAT excludes duplicate strains from analysis and the 62 unique strains [including outlier strains] contained 44 polymorphic sites that had only two alleles, with a minor allele frequency of at least 10%. The influence of recombination on each nucleotide substitution was twice that of mutation \([r/m \text{ ratio}=2.02, \rho=0.006, \theta=0.003]\). The ratio of recombination to mutation for the SA populations that caused SAB and nasal colonisation, with and without the outlier strains is shown in Table 5.4.

<table>
<thead>
<tr>
<th>test</th>
<th>total</th>
<th>blood</th>
<th>nasal&lt;sub&gt;outlier&lt;/sub&gt;</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPT&lt;sup&gt;1&lt;/sup&gt;</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>PHI test statistic</td>
<td>3.1 x 10^-6</td>
<td>2.67 x 10^-4</td>
<td>2.8 x 10^-6</td>
<td>2.4 x 10^-4</td>
</tr>
<tr>
<td>PHI test (P) value</td>
<td>0.001</td>
<td>0.000</td>
<td>0.046</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<sup>1</sup>LPT is the likelihood permutation test
<sup>2</sup>nasal<sub>outlier</sub> = the nasal group including the outlier strains.

Table 5.3 The level of significance for recombination tests performed on nucleotide alignments of the total Auckland sample of SA, and the groups of SA isolated from blood or nasal swab.

Table 5.4 The number of polymorphic sites and estimates of recombination rate \([\rho]\), mutation rate \([\theta]\), and recombination to mutation ratio \([r/m]\) for populations of SA in Auckland from cases of SAB [blood] and healthy nasal carriers [nasal].

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>blood</th>
<th>nasal&lt;sup&gt;1&lt;/sup&gt;</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>polymorphic sites</td>
<td>44</td>
<td>45</td>
<td>41</td>
<td>43</td>
</tr>
<tr>
<td>(\rho)</td>
<td>0.00594</td>
<td>0.00906</td>
<td>0.00375</td>
<td>0.00344</td>
</tr>
<tr>
<td>(\theta)</td>
<td>0.00293</td>
<td>0.00329</td>
<td>0.00307</td>
<td>0.00326</td>
</tr>
<tr>
<td>(r/m)</td>
<td>2.0273</td>
<td>2.7538</td>
<td>1.2215</td>
<td>1.0552</td>
</tr>
</tbody>
</table>

<sup>1</sup>the nasal population including the outlier strains.

In order to examine whether the rates of recombination and mutation in the Auckland SA population were similar to the global SA population, analysis was performed on 30 simulated samples of 62 unique STs, drawn at random from the MLST database [http://saureus.mlst.net/]. The value of \(\rho\) for the Auckland SA population [0.0059] was below the lower confidence interval of the mean value of \(\rho\) obtained from 30 simulated SA populations [mean \(\rho=0.0090, 95\% \text{ CI } 0.0067-0.0114\)]. However, the value of \(\theta\) for the Auckland SA population [0.0029] did not differ significantly from the mean value obtained from 30 simulated SA populations [mean \(\theta=0.0030, 95\% \text{ CI } 0.0025-0.0035\)]. Likewise the \(r/m\) ratio of the Auckland SA population [2.0273] was within the confidence intervals of the mean \(r/m\) value obtained from 30 simulated SA populations [mean \(r/m=3.0885, 95\% \text{ CI } 1.9578-4.2192\)].
5.4.1.9 Recombination in the nucleotide sequence alignment of SA has occurred between housekeeping genes

The analyses above confirmed the presence of recombination in concatenated housekeeping gene sequences; and the same analysis was conducted for each of the seven individual housekeeping genes. The LDHAT likelihood permutation test and the PHI test did not detect recombination in any of the individual housekeeping genes. GARD identified three significant breakpoints in the concatenated sequences, at sites 459, 905 and 1890. The Kishino-Hasegawa topological incongruence test [KH test] compared the topology of neighbour-joining [NJ] trees formed from the alignment upstream and downstream of each breakpoint \(^{142}\). The KH test indicated that NJ trees for each alignment, segregated at each breakpoint, were significantly different \(P<0.001\). The breakpoints at nucleotide 459 and 905 were at either end of the \(aro\) gene. Given that recombination was not detected within the \(aro\) gene, it is likely that the recombination event[s] detected have occurred between MLST loci. The breakpoint at position 1890 is in the vicinity of the \(pta\) gene, but recombination was not detected when only the \(pta\) gene sequences were analysed.

5.4.1.10 SA has lower rates of recombination than Streptococcus pneumoniae and Neisseria meningitidis.

The ratio of recombination to mutation \([r/m]\) was calculated for 100 SA sequences drawn at random from the MLST database [http://saureus.mlst.net/\(^\text{1}\)] and compared with random samples of 100 sequences from bacteria subject to high rates of recombination, \(S.\ pneumoniae\) [http://spneumoniae.mlst.net/\(^\text{1}\)] \(N.\ meningitidis\) [http://pubmlst.org/neisseria/\(^\text{1}\)] and low rates of recombination, \(Bacillus\ cereus\) [http://pubmlst.org/bcereus/\(^\text{1}\)]. Analysis of a large number of simulated samples is required to determine the statistical significance of differences between organisms; but, due to the computational time required [one LDHAT analysis of 100 \(N.\ meningitidis\) genotypes took more than 4 days using a modern desktop PC], only three iterations were performed to ensure convergence of the results.

Recombination was detected in all of the three simulated samples of SA, \(S.\ pneumoniae\) and \(N.\ meningitidis\) \([p<0.001\) for all three tests of recombination]. The measures of \(\theta\) and \(\rho\) were similar for each of the three simulations for each bacterial species [Table 5.5]. \(S.\ pneumoniae\) has very high rates of recombination and high \(r/m\) ratios; \(N.\ meningitidis\) has high rates of recombination and mutation, but the \(r/m\) ratios are comparable to those of SA. The \(r/m\) ratios for SA are higher than those of \(B.\ cereus\), although the values of \(\rho\) were similar.
Table 5.5 Recombination [ρ] and mutation [θ] rates and ratio [r/m] for simulated samples of SA, S. pneumoniae [PN], N. meningitidis [NM] and B. cereus [BC].

<table>
<thead>
<tr>
<th></th>
<th>SA1</th>
<th>SA2</th>
<th>SA3</th>
<th>PN1</th>
<th>PN2</th>
<th>PN3</th>
<th>NM1</th>
<th>NM2</th>
<th>NM3</th>
<th>BC1</th>
<th>BC2</th>
<th>BC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ss</td>
<td>32</td>
<td>43</td>
<td>42</td>
<td>43</td>
<td>52</td>
<td>44</td>
<td>275</td>
<td>267</td>
<td>261</td>
<td>194</td>
<td>197</td>
<td>204</td>
</tr>
<tr>
<td>ρ</td>
<td>0.010</td>
<td>0.009</td>
<td>0.011</td>
<td>0.189</td>
<td>0.151</td>
<td>0.239</td>
<td>0.093</td>
<td>0.058</td>
<td>0.073</td>
<td>0.013</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>θ</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
<td>0.004</td>
<td>0.003</td>
<td>0.016</td>
<td>0.016</td>
<td>0.015</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.014</td>
</tr>
<tr>
<td>r/m</td>
<td>5.18</td>
<td>3.49</td>
<td>4.24</td>
<td>62.61</td>
<td>41.42</td>
<td>77.17</td>
<td>5.76</td>
<td>3.67</td>
<td>4.76</td>
<td>0.98</td>
<td>0.56</td>
<td>0.60</td>
</tr>
</tbody>
</table>

*ss = segregating sites*

5.4.1.11 The sequence types of SA in Auckland

The sequence type [ST] designation is a widely accepted nomenclature system for labelling SA strains and was used to label the Auckland SA strains for phylogenetic analysis. A sequence type [ST] is defined by the allelic designation of seven housekeeping genes. STs are grouped into clonal complexes [CC] defined by STs that share 5 or more alleles. The ST with the most daughters is defined as the ancestor and the CC takes on that ST’s number [Section 5.1.5.1]. The STs and CCs of the Auckland SA population are shown in Figure 5.10A, 5.10B.
Figure 5.10 [A] The sequence types [ST, top] and [B] clonal complexes [CC, below] of 244 SA isolates in Auckland. The number indicates the name of the ST or CC and the percentage of the total is also shown. A number of the labels have been removed from the ST chart; these STs are as follows [clockwise]: 101, 88, 34, 22, 20, 789, 217, 12, 9, 762, 587, 432, 368, 239, 199, 120, 72, 54, 25, and 7.
62 different STs were identified in Auckland; remarkably, 34/244 [14%] of the STs were novel; they did not exist in the MLST database [http://saureus.mlst.net/] and have not been described elsewhere in the world. The commonest sequence type [ST] found in Auckland was ST5 [35/244, 14%] followed by ST1 [29/244, 12%] [Figure 5.10A]. The three most dominant clonal complexes, CC1, CC5 and CC30 include almost two-thirds [151, 62%] of the isolates [Figure 5.10B]. Novel STs are labelled ST11111 to ST11138.

5.4.1.12 The Auckland SA population lacks genetic diversity compared with outlier strains

The degree of genetic diversity of the Auckland SA population was first estimated using a neighbour-joining tree. This tree was formed from a pairwise distance [PWD] matrix for all isolates using SPLITSTREE 139. Almost all of the SA strains form a closely related population [darker grey circle in Figure 5.11] of limited diversity compared with the four outlier strains obtained from healthy nasal carriers, which form an outgroup from the main SA population [lighter grey circle in Figure 5.11]. The outlier strains, confirmed to be SA, were compared with the global database [http://saureus.mlst.net/] and were unique. However, similar strains have been reported, one from a nasal carrier in Cambodia and several from remote aboriginal communities in north of Australia 120, 170. The four outlier strains in Auckland clearly share a common ancestor with clonal complex [CC] 75 strains 170.

Figure 5.11 Phylogenetic relationship of Auckland SA strains. A neighbour-joining tree was reconstructed, using uncorrected P distance, of 244 Auckland SA strains. The small dark grey circle on the left encompasses the diversity of all of the isolates from people with bloodstream infection and the majority of carriage strains. The larger grey circle encompasses the distance from the main SA population to four of the strains isolated from healthy carriers. The branch to the right indicates the position on the tree of S. epidermidis ATCC12228 as a point of reference. The scale bar represents uncorrected genetic distance.
5.4.1.13 The ancestry of the Auckland SA sample: analysis of ST alleles

eBURST was used to display the phylogeny of SA STs in Auckland. eBURST uses allelic designations to link STs that share five or six common loci into clonal complexes (CC), which contain double locus variants (DLV) and single locus variants (SLV). One of the strengths of eBURST is that an allele may alter by recombination or mutation at two loci yet the evolutionary relationship will remain evident. Figure 5.12 shows the eBURST forest created from the Auckland STs. An unrooted tree linking SLVs and DLVs represents each CC. The branch length carries no specific weight in each tree. As an example, CC8 includes four STs that share a recent common ancestor; two of these [ST239 and ST11118] are SLVs of ST8; ST368 is a SLV of ST239 and a DLV of ST8.

The Auckland SA forest consists of nine distinct groups, and the two largest groups contain more than one tree. Each group consists of CCs linked by triple locus variants, that is, STs that are identical at four of seven loci. Six of the ten most common STs in Auckland [ST5, ST1, ST188, ST15, ST8 and ST97] and their CCs, form one group. However, eBURST is not able to derive the relationship between CCs and groups further.

Figure 5.12 The phylogenetic forest of Auckland SA sequence types, created from allelic numbers of seven housekeeping loci using goeBURST V1.2.1. Each box contains the clonal complex(es) that belong to that group.
In addition to the trees shown in figure 6, five singletons [ST20, ST101, ST11128, ST11129 and ST11139] are not included in the figure. These strains, which are not closely related to the remainder of the STs, are of particular interest. ST11128 and ST11129 represent the four CC75 outlier strains mentioned above. ST11139 is a SLV of ST643 in the MLST database [http://saureus.mlst.net/]. Only one example of ST643 has been entered into the database; this was an isolate that caused SAB in a woman in Brazil. The database contains several examples of ST20 and ST101 from all over the world, and each forms a different, small clonal complex after eBURST analysis of the global database [161]. eBURST links novel SA strains identified in Auckland [indicated by the five digit STs in Figure 5.12] with known core STs [e.g. ST1 in Figure 5.12]; this provides indication of local evolution of novel STs from existing STs in Auckland.

5.4.1.14 The ancestry of the Auckland SA sample: phylogeny derived from nucleotide sequences

The phylogenetic reconstruction above, based solely on allelic designations, was unable to define the evolutionary relationship between clonal complexes or groups. Thus, the concatenated nucleotide sequences were also analysed. The phylogenetic tree, shown in Figure 5.13, was reconstructed by maximum likelihood using TREEPUZZLE [127]. Tree puzzle derives the maximum likelihood for a quartet of strains selected at random and then uses these results to derive the most likely tree. For the SA strains a good tree representation was obtained; only 3.1% of the quartets were partially resolved and 1.3% of the quartets were not resolved [triangular insert in Figure 5.13]. Cooper and Feil have previously used evolutionary relationships revealed by phylogenetic analyses to divide SA into three groups [107, 119, 171]. The Auckland SA phylogeny is consistent with this depiction with three exceptions; however, the phylogeny reconstructed by Cooper and Feil is likely to be more accurate – it is derived nucleotide sequences from 37 genes. ST45 was placed in group 1b rather than group 1a; ST22 was placed in group 2 rather than group 1a; and the outlier CC75 strains branched off the hypothetical ancestor by a deep evolutionary branch. The novel strains were distributed throughout the tree, indicating that no particular genetic background is more likely to give rise to a new strain than another.

CLONALFRAME uses Bayesian inference to derive a tree from concatenated nucleotide sequences. It has the advantage of recognising recombination events and reconstructing a tree using the sequence data that is free of recombination; thus preserving the relationship evident in an evolutionary model that includes point mutation only. Figure 5.14 is a tree reconstructed by CLONALFRAME, displayed as a network. Again, the novel Auckland strains occur throughout the network. The CLONALFRAME phylogeny presented as a tree [not shown] is in agreement with the maximum likelihood phylogenetic reconstruction, indicating that the influence of recombination on phylogenetic reconstruction of SA is minimal. CLONALFRAME predicts that one novel strain, ST11122, is an ancestor to a known sequence type, ST25. eBURST analysis showed that CC25 consisted of these two strains [Figure 5.12], but was not able to provide further insight into the evolutionary relationship nor indicate coalescence between ST25 and ST97, which share three identical loci.
Figure 5.13 The phylogeny of Auckland SA strains. The tree was created using TREEPUZZLE 5.2 using 100,000 puzzling steps and the HKY85+G nucleotide substitution model. Branch support [%] is shown as is the number of unresolved quartets in the triangular insert. Novel Auckland STs are shown in grey and the arrow indicates the CC75 STs.
Figure 5.14 Phylogenetic tree of 240 SA strains from Auckland viewed as a network. For display purposes the branch length has not been maintained, but the size of each node is proportional to the number of isolates within the sample. The tree was reconstructed using CLONALFRAME. Strains that have evolved in Auckland are shown in red, the hypothetical ancestral strain is shown in black and CC75 strains have been circled.

5.4.1.15 The ancestry of the Auckland SA sample: the ancestral source of nucleotide polymorphism

STRUCTURE uses a Bayesian method to group recombining populations into their ancestral populations. The presence of linkage between polymorphisms in each sequence is used to derive the ancestral source of that polymorphism. The proportion of polymorphisms that come from each ancestral source determines the ancestry of each strain. The presence of linked polymorphisms is often associated with population structure, such as that which occurs through independent evolution in geographically isolated populations.
Figure 5.15: The ancestral populations defined by STRUCTURE. [A] The proportion of ancestry for individual SA isolates, sorted by the source of the isolate; [B] individual SA isolates sorted by main ancestral group; and [C] the phylogenetic network of SA STs coloured according ancestral group. The colours match in each part of the figure, except in the network group X strains have been coloured pink.
STRUCTURE identified six ancestral groups for the Auckland sample of SA [Figure 5.15A]. Five of the six groups, group A through group E, are homogenous; more than 67% of the genetic background for each isolate originates from one ancestral population. Group E is small and distinct; it contains the CC75 outlier strains isolated from people with asymptomatic nasal colonisation. Isolates in the final group, group X, contain no more than a third of their genetic background from any single genetic background. Indeed, most isolates in group X have three sources of ancestry derived from ancestors of group X [shown in white on Figure 5.15B], group A and group B. The ancestral groups identified by STRUCTURE mirror the phylogenetic relationships within the Auckland SA population, indicated by the phylogenetic network coloured according to ancestral population defined by STRUCTURE in Figure 6.9C. The pink STs [ST9, ST78, ST88, ST188 and ST432] in Figure 5.15C derive ancestry from several groups, suggesting that recombination has been important in the evolutionary history of these lineages. In contrast, a large number of strains belong to ancestral groups B and C, which display almost no evidence of prior recombination.

5.4.1.16 Evidence of SA evolution in New Zealand

The phylogenetic analysis indicated that, in Auckland, new strains of SA originate in a diverse range of genetic backgrounds. A high proportion of the STs identified in Auckland are novel [34/244, 14%]; and have not been identified previously. 16/34 [47%] of the novel STs were isolated from people with SAB, the remainder from nasal colonisation. Each novel ST was compared to the most closely related known ST [the putative parent ST] to determine whether the new allele[s] arose by mutation or recombination and if the change had caused a change in the amino acid sequence [non-synonymous] or not [synonymous] [Table 5.6]. For this analysis mutation was defined as a difference of a single nucleotide compared with the parent ST and recombination was defined as two, or more, nucleotide changes compared to the parent 160. For example, methicillin-resistant SA strain ST11123 was confirmed to be a Western Samoan phage pattern MRSA [WSPP MRSA] by pulsed field gel electrophoresis [performed by Helen Heffernen, Environmental Science and Research, Wellington]. However, the aro gene allele was not allele “2” as expected but was allele “1”. These alleles differed by five nucleotides, of which, G212A led to a change in the amino acid sequence: G71D. The high rate of recombination detected in these novel isolates [19/34, 56%] is at odds with previous estimates that have used the same method, however previous estimates measured recombination in all single locus variants and not just novel isolates. This discrepancy is not caused by sequencing error, because all isolates were sequenced in duplicate and any discrepancies between pairs of isolates were resolved by re-sequencing.
Each new strain was isolated from a different person and was compared to the most closely related known ST to identify variant loci, shaded in grey. The loci shaded dark-grey were found in new STs isolated from more than one individual.

New loci that arose by recombination

NS [non-synonymous] denotes a change in the amino acid sequence; S [synonymous] denotes no change in the amino acid sequence.

Table 5.6 Novel sequence types [ST] in the Auckland population of SA.
5.4.1.17 The influence of selection on housekeeping genes of SA

The changes in the nucleotide sequence of two thirds of the new strains [23/34, 68%] resulted in a change in the amino acid sequence. Single-Likelihood Ancestor Counting [SLAC] detected 20 polymorphic sites under negative selection and no sites under positive selection. The global ratio of non-synonymous nucleotide changes, dN [that gives rise to an altered amino acid sequence], compared to the rate of synonymous change, dS [that does not give rise to an altered amino acid sequence], was 0.093. This indicated that, on average, the rate of synonymous mutations is 11 times greater than the rate of non-synonymous mutations. Thus, non-synonymous mutations in the housekeeping genes of SA, which encode for enzymes with metabolic functions, do not often persist.

5.4.1.18 Migration: evidence of transmission of novel sequence types

There were several examples of transmission of novel SA strains. On three occasions unique alleles were identified in SA strains from more than one person. ST11126 has a novel tpi allele; this strain was isolated from three people, who lived together in the same household. ST11119 has two point mutations in the yqi allele, and was isolated from two unrelated people with SAB. A further unrelated person had SAB caused by a ST with the same yqi allele, but this ST [ST11120] also has an additional point mutation at the tpi allele. ST11119 and ST11120 form a CC with ST12 as shown in Figure 5.12.

ST11128 was isolated from three unrelated people and has novel arc and gmk genes. The novel arc gene is also shared with ST11129, isolated from another unrelated person. These two STs were the outlier CC75 strains. Finally, two unrelated people had SAB caused by ST11131. All of these alleles [6, 5, 6, 48, 7, 14, 5] are known, but the gmk allele has not been found in this genetic background before. These strains were single locus variants of ST121.

5.4.1.19 Comparison of the sequence types that cause nasal colonisation and bloodstream infection

The Auckland SA population of 244 strains were derived from two distinct environments. The isolates that caused SAB [n=150], and the isolates from healthy nasal carriers [n=94] were remarkably similar, with the exception of the four outlier strains in the nasal group. Figure 5.11 shows there are few differences between the population of SA that cause SAB and the population of SA that cause nasal colonisation. The ten largest STs contain 188 [77%] of the SA isolates. There is no common ST isolated exclusively from people with SAB [the ‘blood’ group] or from nasal carriers [the ‘nasal’ group]; although ST1 isolates are more likely to cause SAB [25/150, 17%] than nasal colonisation [4/94, 4%] [Fisher’s exact test, p=0.0038] [Figure 5.16]. ST188 is a member of CC1 [it shares five of its seven alleles with ST1] but has no predisposition to cause SAB. 13/94 [14%] of the nasal isolates were ST188, compared with 12/150 [8%] of the blood isolates; this difference was not statistically significant [Fisher’s exact test, p=0.1924].
5.4.1.20 Testing the null hypothesis of no difference between the blood and nasal groups of SA

Even though ST1 causes more SAB than nasal colonisation, a closely related strain, ST 188, does not. Obviously there are limitations in this type of comparison, based solely on the ST, which does not take into account the fact that many of the STs are closely related. A comparison of the SA strains isolated from cases of SAB or nasal carriers [not including CC75 outlier strains] was performed using clusters of OTUs derived from the PWD of each subpopulation. Thirteen OTU_{0.0049}s [SA strains grouped together into an operational taxonomic unit with PWD of less than 0.0049] were identified in the total sample; 11/13 (85%) OTU_{0.0049}s are shared between the sample of SA from cases of SAB and the sample of SA from nasal carriers [Figure 5.17]. The blood group contained two OTU_{0.0049}s not present in the nasal group.

* bloodstream infection versus nasal carriage, Fisher’s exact test, p=.0038.

Figure 5.16 The percentage of the nine most common sequence types [ST] and the new strains, which caused blood or nasal infection in 244 people with SA infection in Auckland.
The parsimony test was implemented in MOTHUR to screen for differences between two populations. A parsimony score comparing the blood and nasal isolates within the same neighbour-joining tree was calculated and then the taxa were randomly shuffled to determine how often that same score, or lower, could arise by chance. The score for the blood and nasal subpopulations was 37; the lowest score generated from 1000 random permutations was 62 indicating a statistically significant difference between the two populations ($P<0.001$). While a phylogenetic difference could be detected, further analysis was required to determine whether there were substantial differences between the two groups.

The LibShuff method, also implemented in MOTHUR, was used to detect differences in the genetic diversity between the blood and nasal groups. LibShuff calculated a statistic to represent the difference between each group’s collectors curve obtained by comparison of the proportion of coverage of the population versus genetic distance [Section 5.1.8.1]. 1000 random shuffles of the strains were performed to identify how often the difference between the curves [or greater] was identified by random chance. LibShuff failed to find a significant difference between the blood and nasal isolates ($P=0.09$).

The genotypes of the populations of SA obtained from people with SAB and those with nasal colonisation overlap when plotted in two dimensions using multi-dimensional scaling [Figure 5.18]. Any differences are minor as the populations are not significantly different [ANOSIM, $R=0.031$, $P=0.062$; PERMANOVA, pseudo $F=2.28$, $P=0.092$].
Finally, STRUCTURE was used to look for differences in the ancestral populations between SA isolated from people with SAB or nasal colonisation. STRUCTURE identified six populations in the total Auckland SA population [Figure 5.19]. The CC75 strains, form a discrete ancestral population [green bar in Figure 5.19], only present in the nasal group. Otherwise, the remaining five ancestral populations were present in both groups of SA isolates.

5.4.1.21 SA lineage that causes infection in Māori and Pacific people in Auckland

52 of the 240 [22%] strains in the Auckland sample, caused infection in Pacific people and 39 [16%] caused infection in Māori. The remaining 149 [62%] were isolated from people of other ethnic groups. Comparison of the SA strains that caused infection in these three ethnic groups did not find significant differences [PERMANOVA, pseudo F = 2.22, p=0.076] [Figure 5.20].
were no significant differences in the six ancestral populations defined by STRUCTURE between different ethnic groups.

Figure 5.20 2D MDS plot of 240 strains of SA in Auckland by ethnic group.

For ease of comparison and for display purposes, Māori and Pacific people were grouped to examine the proportion of ethnicity of the people infected with SA by sequence type [ST]. In Figure 5.21, STs that cause infection in Māori or Pacific people are coloured red; STs that cause infection in people of other ethnic groups are coloured in blue. A grey scale is used for other STs, with lighter grey indicating fewer infections in Māori or Pacific people. 91/240 [38%] of the people infected with SA in the Auckland sample were Māori or Pacific people.
There are two main observations. Firstly, ST97 contains six isolates, five of these [83%] were from Māori [n=4] or Pacific [n=1] people. Secondly, all of the seven strains in CC121 [ST120, n=1; ST121, n=4; novel strain ST11131, n=2] were only found in two Māori and five Pacific people. There are no branches of the phylogeny that contain only strains from people of other ethnic groups; although 8/9 [89%] of ST508 isolates and novel variants of ST508 were isolated from non-Māori and non-Pacific people.

Previous reports and the analysis presented in Chapter 2 suggested that a particular ST30 strain of methicillin resistant SA [MRSA] known as the Western Samoan phage pattern of MRSA [WSPP MRSA, also known as the Southwest Pacific clone of MRSA] causes more infections in Māori and Pacific people living in Auckland [36, 37]. However, CC30 infections [caused by methicillin susceptible or resistant isolates] in the Auckland sample were not more common in Māori and Pacific people [Figure 6.15]. It is likely that MLST lacks the resolution to detect relationships between SA strain and host ethnicity.

**5.4.1.22 Further resolution of STs within CC30 by PFGE did not provide additional correlation with ethnic group or methicillin resistance**

In order to determine whether another typing method was able to correlate SA isolates with phenotypic or host characteristics, pulsed field gel electrophoresis [PFGE] was performed on all
of the CC30 strains by Helen Heffernen [Institute of Environmental Science and Research, Wellington, the New Zealand reference laboratory for staphylococcal typing]. Only one subtype of CC30 strains [all ST30] is methicillin-resistant; all of the five MRSA isolates [including one novel strain] had an identical PFGE pattern to the WSPP reference strain. These isolates were obtained from two Māori people, one Pacific person and two people from other ethnic groups. However, 5/13 [38%] methicillin-susceptible CC30 strains were WSPP MSSA; the PFGE pattern did not correlate with methicillin-resistance.

The PFGE types of methicillin-susceptible isolates are shown in Figure 5.22. Five isolates were very closely related to the WSPP MRSA reference strain, one was identical. PFGE groups of strains with more than 90% homology did not correlate with the strain’s ST. The first group contained two ST30 isolates [alleles: 2, 2, 2, 2, 6, 3, 2] and one isolate of ST39 [alleles: 2, 2, 2, 2, 2, 2, 2]. ST39 is a DLV of ST30, the *pta* and *tpi* nucleotide sequences each differs by a single point mutation. There was no association between PFGE group and host ethnicity.

* Methicillin susceptible isolates that share more than 90% homology with WSPP MRSA

Figure 5.22 Dendrogram of homology of PFGE of *SmaI* digested genomic DNA of methicillin susceptible CC30 isolates [performed by Helen Heffernen, ESR]. The ST and ethnicity of the host of each isolate is listed; isolates were grouped if they shared more than 90% homology by PFGE type. The reference strain of ST30 WSPP MRSA is also displayed.
5.4.1.23 The analysis of genes in addition to the conventional housekeeping genes does not add to the ability of MLST to differentiate strains of SA

Cooper and Feil reconstructed a robust phylogeny from the concatenated sequences of 37 gene fragments from 30 strains of SA \(^{171}\). Their phylogeny utilised housekeeping genes [including the seven conventional housekeeping genes], cellular envelope genes, genes involved in informational pathways, orphan genes and genes of unknown function. It is possible that analysis of gene segments under neutral or positive selection, in addition to the concatenated sequences of seven conventional housekeeping genes, might allow further discrimination of SA. A previous study that examined genes under negative selection in addition to the MLST housekeeping genes demonstrated that additional genes did not provide further resolution compared with conventional MLST alone \(^{173}\). To test this further, the ss/7 genes from 94 strains of SA, isolated from cases of SAB at Auckland City Hospital, were examined. Staphylococcal superantigen-like protein 7 [SSL7] is a putative virulence factor, which binds complement factor C5 and IgA and prevents formation of C5a and the membrane attack complex [Section 6.1.14].

The ss/7 gene was found in all 94 isolates of SA. An alignment of 603 nucleotides, which covered the entire open reading frame of the gene, was analysed. There were 135/603 [PWD 0.22] polymorphic sites in the ss/7 alignment, which was greater than the degree of polymorphism in the alignment of concatenated MLST loci for the same 94 isolates [90/3198 [PWD 0.03]. The presence of recombination within the nucleotide alignment was not confirmed by agreement between the PHI test [recombination detected, PHI = 6.98 x 10^{-5}, P<0.001] and the LDHAT likelihood permutation test [no recombination detected, P=0.25]. Two recombination breakpoints were identified, at nucleotides 229 and 371 [GARD, KH testing of tree topology, \(P<0.001\)]. The estimated mutation rate, \(\theta\), was 0.031 and the recombination rate, \(\rho\), was 0.008. The ratio of recombination to mutation, \(r/m\), was 2.67; the same as the ratio for the concatenated housekeeping gene sequences of isolates from 150 cases of SAB.

The concatenated housekeeping gene sequences are under strong purifying selection; in contrast the ss/7 gene alignment is under neutral selection, the global dN/dS ratio was 1.011. This indicates that changes in the nucleotide sequence are as likely to be synonymous [no change in the amino acid sequence] as non-synonymous. Five codons are under significant negative selection pressure [amino acids 36, 37, 38, 67, and 109; numbered from the mature +1 residue], which suggests that the amino acid residues encoded by these codons have important functions. Only codon 151 is under significant positive selection.

5.4.1.24 The functional importance of ss/7 codons under selection pressure

The codons under selection pressure were mapped onto the amino acid alignment of the ss/7 gene [Figure 5.23]. All of the codons under negative selection pressure represent sites that are known to be important for binding of SSL7 to IgA \(^{174,175}\). The site under positive selection [codon 151] is important for C5 binding \(^{174}\).
5.4.1.25 Co-evolution of ssl7 with SA housekeeping genes

The ssl7 gene is found within pathogenicity island 2 [SAPIs2], a mobile genetic element that contains a variable assortment of 11 ssl genes and two regulator genes hsDs and hsDm₁⁷⁶,₁⁷⁷. A maximum likelihood phylogenetic tree was reconstructed using ssl7 gene sequences from 94 SA strains isolated from cases of SAB at Auckland City Hospital. In almost every circumstance the taxa grouped in agreement with the phylogeny reconstructed from the concatenated housekeeping gene sequences [Figure 5.24]. The co-evolution of ssl7 with the SA housekeeping genes indicates that ssl7 is not able to aid discrimination of SA strains.

One ST45 strain [A94] had a ssl7 gene sequence that was closely related to ST101 strains and unrelated to other ST45 strains. This is likely to represent a recombination event involving SAPIs2. Two ST20 isolates [A55 and A28] and one ST9 isolate [A64] were grouped with CC15 strains of SA, yet neither ST9 nor ST20 belong to CC15. The phylogenetic tree in Figure 5.13 shows that these group 2 STs were closely related; and it is likely that they do share a common [or very similar] SSL7 allele.
Figure 5.24 The phylogeny of 94 SA strains isolated from cases of SA bacteraemia at Auckland City Hospital reconstructed from *ssl7* gene sequences. The clonal complexes [CC] or sequence types [ST] are also shown; the taxa highlighted in pink were "misplaced" in the *ssl7* phylogeny.
5.4.2 Comparison of the genetic structure of SA populations in Auckland, Samoa, Fiji and Tonga

An increased prevalence of infections caused by ST30 MRSA was identified in Māori and Pacific people living in Auckland in the 1990’s\textsuperscript{36,37}; however, there was no increased prevalence in infections caused by ST30 SA [MRSA and MSSA] in the large sample of SA from people in Auckland described in this thesis. The proportion of Māori and Pacific people with infections caused by CC30 STs [ST30, ST34 and ST39] was similar to the proportion of Māori and Pacific people in the total sample [Figure 5.21].

The finding that all of the CC121 strains were isolated from Māori or Pacific people in Auckland warrants further examination on a regional level to determine whether CC121 strains are prevalent in Pacific populations. If different sub-clones of SA cause different rates of infection in different ethnic groups, then one would expect similarity between populations of SA that cause infection in Pacific people living in Auckland and people living in Pacific Island nations. However, unless modern human travel has completely mixed the populations of SA, one would also expect there to be numerous examples of STs that cause infection in one country but not in another.

Auckland is home to the largest community of Pacific people in the world. In 2007, 7% of the people who received permanent residence status in New Zealand came from Fiji, 5% came from Samoa and 2% came from Tonga\textsuperscript{176}. The number of migrants from these three countries was greater than the number of migrants from other Pacific nations. Thus, samples of SA were obtained from Fiji, Samoa and Tonga and compared to the genetic population structure of SA in Auckland, in order to understand the genetic population structure of SA in our region.

5.4.2.1 The sample of SA strains from Samoa

The SA strains from Samoa were obtained from a study of the prevalence, susceptibility profile and genotypes of MRSA that cause wound infection in Samoa [Chapter 3]. SA was isolated from the wounds of 187/399 [47%] people who had wound infection. 66 [35%] of the SA infected participants were female and their median age was 24 years. Multi-locus sequence typing [MLST] was performed on 96 strains selected at random from the 187 SA strains. The random sample included 26/34 MRSA strains isolated from study participants.

5.4.2.2 The sample of SA strains from Fiji

After ethical approval was obtained from the Fijian Ministry of Health, SA isolates from clinical specimens processed by the microbiology laboratory of the Colonial War Memorial Hospital [Suva, Fiji], were collected. This is the only public laboratory in Fiji’s largest city, and it also receives a small number of clinical specimens from other parts of Fiji, including the hospital at Lautoka and a smaller healthcare centres. Consecutive strains of SA, isolated from different hosts, were collected between January 2008 and August 2008. The isolates of SA came from a number of clinical specimens, but the majority [163/205, 80%] were associated with wound
infections. The median age of the study population was 35 years [interquartile range 14.5 to 54 years]. Access to the clinical records was not possible and only limited information was available from the specimen request form. In particular, the patients’ names were not available, and it was not possible to determine whether each infected person was of Fijian or Indian ethnicity. MLST was performed on 109 isolates randomly selected from the total sample of 205 isolates. The median age of the random sample population was 33 years [interquartile range 13 to 52 years] and 85/109 [78%] of the strains were isolated from wound swabs. None of the isolates were MRSA.

5.4.2.3 The sample of SA strains from Tonga

In Tonga a small sample was obtained from the hospital microbiologist, Mary Fakahau. These 18 isolates of SA were obtained from patients in Vaiola Hospital [Nuku’a’lofa, Tonga]. Two isolates were from blood cultures, three were isolated from urine culture, one was isolated from sputum and the remaining 12 were isolated from pus swabs taken from a range of infected sites. All of these strains were collected between September 2007 and March 2008, and all were believed to be clinically significant. The only demographic information available was the information available on the laboratory request form. All patients were Tongan and their median age was 19 [6months – 67 years]. None of these isolates were MRSA.

5.4.2.4 The SA samples from Samoa and Fiji were representative of the total SA population in each country

Rarefaction curves indicate that the genetic diversity of SA from Samoa and Fiji had been adequately captured. When the strains were grouped into OTU_{0.0049}s the number of OTUs identified did not continue to rise as each sample reached its maximum number. Even the small sample of strains from Tonga captured the genetic diversity to a reasonable degree, but the rarefaction curve for OTU_{unique} continued to rise relatively steeply, indicating that this sample was not adequate to explore the evolution of new strains in Tonga [Figure 5.25A]. The rarefaction curve for the total sample of SA from Pacific Island nations is shown in Figure 5.25B.

Efron’s calculation indicated that doubling the sample from Tonga would have identified a further seven unique strains, but would not have increased the number of OTU_{0.0049}s. In Samoa and Fiji, increasing the sample by 10% would have identified one further unique strain, but no further OTU_{0.0049}s. For the combined Pacific sample of 223 SA isolates, a further 3 unique strains would have been identified by increasing the sample by 10%.
Figure 5.25 Rarefaction curves for the SA sample obtained [A] from 18 people in Tonga; and [B] from the total Pacific sample of 223 strains.
5.4.2.5 The STs of SA that cause infection in Samoa and Fiji were similar

29 unique STs were identified in the Samoan sample of 96 SA isolates from wound swab culture. In the Fijian sample of 109 isolates, 26 unique STs were identified; and in the Tongan sample of 18 isolates, 9 unique STs were identified. In Samoa, the ten largest STs contained 69 [72%] of the isolates; in Fiji the ten largest STs contained 85 [78%] of the isolates [Figure 5.26A and B].

---

Figure 5.26 The sequence types [ST] identified in [A] the Samoan sample of 96 SA isolates; and [B] the Fijian sample of 109 isolates. The labels of STs with only one isolate have been removed for display purposes.
The degree of sharing of STs found in Samoa, Fiji and Tonga is shown in the Venn diagram in Figure 5.27. There are a number of STs unique to Fiji and Samoa, which also share 12 STs. The small sample from Tonga contained only one unique ST and this was ST45, which was isolated several times in Auckland. When the strains were grouped into OTU$_{0.0049}$, there were no OTU$_{0.0049}$s found in Tonga that were not also found in the sample from Samoa; there were 5 OTU$_{0.0049}$s unique to Samoa and 1 OTU$_{0.0049}$ found only in Fiji.

![Venn diagram of shared sequence types of SA in Samoa, Fiji and Tonga.](image)

**Figure 5.27 Venn diagram of shared sequence types of SA in Samoa, Fiji and Tonga.**

5.4.2.6 Differences in abundant STs that cause infection in Fiji, Samoa or Auckland

In Auckland, there was minimal difference between SA strains that caused SAB and nasal carriage, and in the present chapter these samples of SA will be combined for analysis. Some of the ensuing analyses will not include the isolates from Tonga, because the sample was small and the diversity of SA in Tonga was almost entirely included within the Fijian and the Samoan samples.

The sequence types of SA isolated in Auckland, Samoa or Fiji are shown in Figure 5.28. There are a number of STs that were present in small numbers at one site, but not at another, but there are also several interesting differences with common STs. ST1, common in Samoa and Auckland, was not found in Fiji at all. However, a small number of related STs that are members of CC1 [ST573, n=2; novel ST, ST11150, n=3] were found in Fiji, but not at other sites. ST573 is a single locus variant of ST1; it differs by a single point mutation in the *pta* gene.

All of the other common STs found in Auckland were found in the Pacific nations except ST59 which was not identified in Fiji [10, 4% isolates in Auckland and 1, 1% isolate in Samoa were
The proportion of ST30 strains found in Fiji was twice as high as Auckland, but the related single locus variants ST34 and ST39 were not found in Fiji or Samoa.

ST121 strains, found only in Māori and Pacific people in Auckland, are common in Samoa and Fiji. In the Samoan sample 13/96 [14%] of the isolates were ST121; in Fiji 18/109 [17%] isolates were ST121. The novel strain, ST11131 is a member of CC121 found in two Samoan people with SAB in Auckland. ST11131 was also found in Samoa [n=3], Fiji [n=5] and Tonga [n=1] [Table 7 below]. ST508, found only in people of other ethnic groups in Auckland, was isolated once in Samoa.

There were several STs which were common in the Pacific, but not found in Auckland: ST93 [Queensland strain MRSA, found in Samoa]; ST96 [common in Fiji] and ST779 [found in Fiji and Samoa]. The proportion of novel STs [that had not been described before] was similar in each country.
Figure 5.28 The proportion of sequence types of SA found in Auckland [n=244], Samoa [n=96] and Fiji [n=109].
5.4.2.7 The Phylogeny of SA in the Pacific

The list of STs identified in each SA sample from different Pacific countries indicated numerous differences between one location and another, but this simple analysis was not able to take account of the fact that many of the STs were very closely related. Phylogeographical analysis was performed to identify more meaningful geographical differences between related STs.

The maximum likelihood phylogenetic tree of the Pacific sample of SA reconstructed using TREEPUZZLE 5.2 is shown in Figure 5.29. The branches all received a high level of support and the quartets were also well supported, only 3.7% of the quartets were not resolved. The CC75 outlier STs [ST11128 and 11129] found in Auckland and Fiji were included to form an outgroup; and novel STs [denoted by labels 11111-11150] were found throughout the phylogeny. The tree also shows the phylogenetic groups of SA, described by Feil and colleagues 107,119. The new variant ST11142 did not have any single or double locus variants in the MLST database, but was closely related to other group 2 STs.

5.4.2.8 Limited evidence of geographical clustering within the phylogeny of SA in the Pacific

The phylogeny of the 467 Pacific SA strains from Auckland, Samoa, Fiji and Tonga was reconstructed using CLONALFRAME and the dot plot of the phylogeny is displayed as a network in Figure 5.30. The STs found at each location [see also Figure 5.28] are colour coded. All of the abundant STs are found in both Auckland and Pacific Island nations.

One major branch contained strains found only in Fiji or Samoa. The branch contained ST96 [alleles: 12, 1, 1, 15, 11, 1, 40] and ST779 [alleles: 3, 138, 1, 15, 100, 1, 10] which are identical at three loci but differ by only 6 nucleotides in their concatenated sequences. However, the maximum likelihood tree [Figure 5.29] did not place ST96 and ST11146 on the same branch as ST779; on this tree ST779 was equally related to a number of other group 2 STs.
Figure 5.29 The phylogeny of SA STs isolated in Auckland [n=244]; Samoa [n=96] and Fiji [n=109]. The tree was constructed by maximum likelihood using TREEPUZZLE 5.2. The triangular insert indicates the proportion of resolved quartets [the sum of the vertices, 92.8%] from 1,000,000 random quartets of taxa.
The only abundant ST unique to Samoa was ST93; these were all Queensland strain MRSA and probably reflect exposure related to travel between Samoa and Australia. Several STs were unique to the Auckland SA population but in all instances there were closely related STs found in Fiji and Samoa. The large numbers of isolates in the CC121 lineage identified in the Samoan and Fijian SA populations, combine to make CC121 one of the dominant lineages in the Pacific.

Figure 5.30 The phylogeny of 449 isolates of SA obtained from Auckland [n=244]; Samoa [n=96] and Fiji [n=109] constructed using CLONALFRAME. The ST nodes have been colour coded by country; blue, red, and yellow nodes contain isolates identified only in Auckland, Samoa, and Fiji respectively. Green, purple and orange nodes contain isolates identified in Auckland and Fiji, Auckland and Samoa and Fiji and Samoa respectively. Grey nodes contain isolates obtained from Auckland, Samoa and Fiji. For display purposes, the branch lengths have not been maintained.
5.4.2.9 ST93, ST96 and CC121 cause infection only in Māori and/or Pacific people

In the Auckland sample of SA, CC121 strains were isolated only from Māori and Pacific people. ST121 strains were more common in Samoa and Fiji, than in Auckland. The same analysis was extended to include the samples from Samoa and Fiji [Figure 5.31]. This larger SA population contained all of the isolates from Fiji, Samoa and Auckland; and 281/449 [63%] of the isolates were obtained from Māori or Pacific people, living either in Auckland or Fiji or Samoa. Only a few STs contain isolates that were only obtained from people of other ethnic groups and all of these STs contained only one or two isolates. In contrast there were several larger STs that included isolates obtained only from Māori or Pacific people [large red nodes in Figure 23]. Two lineages, which only contained isolates obtained from Māori and Pacific people were identified: the CC121 lineage and the lineage that contained ST96 and ST779. However, as already mentioned, the 50% consensus tree created by maximum likelihood [Figure 5.29] did not group ST96 and ST779 on a discrete evolutionary branch.

Figure 5.31 Phylogenetic network of 449 isolates of SA obtained from Auckland [n=244]; Samoa [n=96] and Fiji [n=109] showing the proportion [%] of isolates isolated from Māori or Pacific people living in Auckland, Samoa or Fiji.
5.4.2.10 Minimal difference in the population structure of SA that infect people living in Pacific Island nations or that infect people living in Auckland

Figure 5.32A shows the Venn diagram of OTU\textsubscript{unique}s [operational taxonomic unit that contains only identical, unique strains and is equivalent to a ST] shared between the samples of SA from Auckland, Samoa, Fiji and Tonga. Only six OTU\textsubscript{unique}s were found at all four sites, and nine OTU\textsubscript{unique}s were shared between the larger samples [Auckland, Samoa and Fiji]. The largest sample, from Auckland contained 41 OTU\textsubscript{unique}s not found at other locations, but most of these were novel strains isolated from only one person.

The Venn diagram for shared OTU\textsubscript{0.0049}s [operational taxonomic unit that contains all similar strains with a PWD of less than 0.0049] is shown in Figure 5.32B. Most of the genetic diversity is contained within the samples from Samoa and Auckland. There were only two unique OTU\textsubscript{0.0049}s in Auckland, four in Samoa, but none in Fiji; all of the nine OTU\textsubscript{0.0049}s in the Fijian sample were found in Auckland and Samoa.
5.4.2.11 No difference in the diversity of the SA population in different Pacific nations

There were more OTU\textsubscript{0.0049}s identified in Samoa than in Fiji or Auckland [Figure 5.32B], which suggests that the diversity of the Samoan sample was greater. However, the diversity of each sample is also dependant on the number of members of each OTU. Simpson’s Index of diversity was calculated from the sum of the number of OTUs multiplied by the number of individuals in each OTU and then divided by the number of isolates in the total population. Samples with greater diversity have lower values. The diversity of each SA sample was not different at any location and Figure 5.33 shows the values of 1-Simpson’s index for OTU\textsubscript{unique} and OTU\textsubscript{0.0049}.
5.4.2.12 Statistically significant differences between the genetic structure of SA populations from different Pacific nations

The Parsimony test was used to screen for a difference between samples of SA from Auckland, Fiji, and Samoa\(^\text{146, 172}\). The parsimony score was calculated for each pair of sample locations within a phylogenetic tree. The taxa were shuffled 1000 times and new parsimony scores were calculated for each permutation. None of the permutations for any of the pairwise comparisons between samples from Auckland, Fiji and Samoa gave a parsimony score as low as the true value and the \( P \) value for all tests was \(<0.001\).

LibShuff identified a significant difference \([P<0.008\) to allow for multiple comparisons\] between pairwise comparisons of the SA populations from Samoa, Fiji and Auckland\(^\text{146}\). The outlier CC75 isolates were not included in the analysis, to avoid overestimating the difference between samples.

5.4.2.13 Significant differences in the genetic structure of SA populations from different Pacific nations were caused by minimal differences between populations

The two dimensional multi-dimensional scaling (MDS) plot of the total Pacific SA population is shown in Figure 5.34. The CC75 isolates were not included on the MDS plot, because their extreme genetic distance from the other strains caused the remaining strains to form a small indistinct bunch. There was no overt separation between the genotypes of SA obtained from each sample location, which formed two main populations, circled in Figure 5.34. The SA populations from Fiji and Tonga are not distinct from the samples from Auckland and Samoa; however, the Auckland and Samoan SA populations contain a small number of more divergent genotypes distinct from the main population.
The SA population from Auckland did not differ greatly from the combined Pacific Island SA population [Fiji, Samoa and Tonga]; ANOSIM [R=0.041] indicated that variability within each population is very similar to the variability in the total Pacific population. The low R value indicates that the null hypothesis might be true, although the permutation test did find a significant difference between the Auckland SA population and the combined Pacific island SA population [ANOSIM, R=0.041, P<0.01; Figure 5.35A]. This finding indicates that the very small difference detected between Pacific and Auckland SA populations was statistically significant due to differences in the small number of more divergent SA strains present in the large sample.

Likewise, comparison of the Samoan population with the combined SA populations from Auckland, Fiji and Tonga yielded similar results [ANOSIM R=0.041, P=0.035; Figure 5.35B]. The SA populations from Fiji and Tonga did not differ significantly from the other SA populations combined [ANOSIM Fiji R=0.027, P=0.08; Tonga R=0.061, P=0.11]. There were no significant differences when populations from individual countries were compared.
5.4.2.14 Novel SA strains in Pacific Island nations with small human populations

The sample of SA obtained to represent the population of SA in Tonga, was too small to allow analysis of unique, novel STs. However, the proportion of novel STs [STs that have not been identified previously] identified in Samoa [9%] and Fiji [16%] was similar to Auckland [14%] [Figure 5.28]. Table 5.7 shows the novel STs found in Samoa, Tonga and Fiji. In the Pacific sample 27/233 [12%] of the strains were novel and 20/27 [74%] arose by recombination. For 24/27 [89%] of the new STs, the change in the nucleotide sequence caused a change in the amino acid sequence when compared to the most closely related ST. One of the new strains [ST11142] had no single or double locus variants in the MLST database. In the phylogeny shown in Figure 21 this strain was placed in group 2. When a neighbour joining tree was used to reconstruct the phylogeny [not shown], ST11142 was most closely related to ST20.
Each new strain was isolated from a different person and was compared to the most closely related known ST to identify variant loci, shaded in grey. The loci shaded dark-grey were found in new STs isolated from more than one individual.

New loci that arose by recombination

NS [non-synonymous] denotes a change in the amino acid sequence; S [synonymous] denotes no change in the amino acid sequence.

ST11142 does not have a single or double locus variant; it was most closely related to ST20 on a NJ tree.

Table 5.7 Novel sequence types [ST] of SA in Samoa, Tonga and Fiji.
5.4.2.15 Migration of novel STs in the Pacific

There was evidence of transmission of new strains between different hosts on several occasions. The most striking example was the CC121 strain ST11131, which has disseminated to all four countries sampled: Auckland, Fiji, Samoa and Tonga. ST11129 was isolated in Fiji and also in Auckland. The genetic variation between this ST and the main population of SA that caused BSI or asymptomatic nasal carriers in Auckland was marked, and was similar to CC75 strains previously identified only in Australia and Cambodia\textsuperscript{120, 170}. There were two further examples of novel strains isolated from different people in Samoa [ST11148] and Fiji [11151].

5.4.3 Migration and evolution of SA in the global SA population

Many pathogens display differences in genetic structure in populations obtained from different locations. Recent studies have demonstrated the relationship between genotypes of \textit{Mycobacterium tuberculosis}, HIV and Dengue fever virus and their country of origin [phylogeography]\textsuperscript{179-181}. This part of the Results places the Pacific SA population into context with the global SA population and compares the phylogeography and evolutionary history of SA throughout the world. The global SA population is represented by 73 STs isolated in Oxford, England, ten years ago; 33 STs isolated from children in China; 20 STs from Mali and the 87 STs isolated from the Pacific and described in this thesis\textsuperscript{107, 112, 119}. The populations of SA selected to represent the strains found worldwide consist of a single representative of each ST found at each location. Thus, the sample encompasses the genetic diversity of SA found at each location, but does not include information about the frequency of each ST at each location. Analyses reliant on this information, such as measures of diversity or rarefaction are not appropriate. Instead, a number of methods were used to explore whether the population structure of SA was similar at each location.

5.4.3.1 Geographical variation and evolution in SA phylogeny

Differences in phylogeography were explored using a phylogenetic network created using CLONALFRAME\textsuperscript{128}. The phylogenetic relationships between SA populations from the Pacific [Auckland, Samoa, Fiji, and Tonga], Oxford, China and Mali are shown in Figure 5.36. The STs that were found only in one of the four continents are colour coded. The network was created from a 50% majority rule tree of the consensus of three duplicate iterations of CLONALFRAME\textsuperscript{128}. The phylogeny indicates that ST152 [yellow diamond in Figure 5.36], found in Mali, branches directly off the hypothetical ancestral sequence [see also Figure 5.42]. A number of STs were found in every population [green in Figure 5.36]. Many of these were among the predominant STs found in each population, for example ST30 strains were the most abundant in Oxford and Fiji; ST121 strains were the most abundant in China and were also abundant in Samoa and Fiji.
A large number of STs are found only in the Pacific SA population [coloured black] or in Oxford SA population [coloured blue]. Although these STs are distributed throughout the phylogeny, there is a concentration of STs found only in Oxford within CC30 [around ST30 and ST39], CC15 [around ST14, ST15 and ST199], CC22 [around ST22] and CC25 [around ST25]. These STs were the most abundant in Oxford 10 years ago, and it is likely that the STs found only in Oxford represent novel STs that have evolved from these predominant clones. Likewise, STs found only in the Pacific SA population are concentrated within CC30 [around ST30 and ST34], CC5 [around ST5], near ST188, and within CC8 [around ST8]. Again, the STs found only in the Pacific represent novel STs that have evolved from abundant STs. In the global population of SA novel STs within CC121 are found in China, Oxford and the Pacific, indicating that micro-evolution of STs within CC121 is occurring in diverse locations.

The STs identified in Mali consisted of pandemic STs, novel STs and ST152. The novel STs in Mali are distributed throughout the phylogeny. However, the situation in China is different. Novel STs in the SA population from China are distributed throughout the phylogeny with the notable exception of novel STs within CC30 and CC1, which are less common in the population of SA from China. Novel STs in the SA population from China are found in the most abundant CCs in China, CC121 and CC59.
5.4.3.2 Differences between SA populations from Pacific island nations and Mali

It is reasonable to assume very little human travel between Pacific Island nations [Fiji, Tonga and Samoa] and Mali. In the absence of migration, it is reasonable to assume that samples of SA collected from Mali would differ from samples of SA collected from Fiji, Samoa or Tonga. In Figure 5.37 the STs found in Mali and the Pacific Island nations are highlighted on the network created from the global SA population. Most branches in the network include STs found in both Mali and the Pacific Island nations, except for branches to CC30 and CC59. In all of the other major branches, STs found only in the Pacific Island nations can be found, and, in almost all circumstances, there is a closely related ST found in the Mali SA population.

Figure 5.37 The phylogeny of SA in the Pacific and Mali superimposed on the phylogeny of the global sample of SA. The white nodes represent STs that were not found in Mali or the Pacific Island nations [Fiji, Samoa and Tonga]. Pink nodes represent STs found in Mali and in Pacific Island nations. Black nodes represent STs that were found only in Pacific island nations and not in Mali; yellow nodes represent STs that were found only in Mali and not in Pacific Island nations.

5.4.3.3 Geographically diverse SA populations are similar

Further analysis was required to determine if the differences shown in Figure 27 were meaningful at a broader evolutionary scale. STRUCTURE was used to detect differences in the ancestral populations of SA in the global sample of SA. STRUCTURE identified seven sources of ancestry in the global sample [Figure 5.38A]. 33 STs [14%] were assigned a new group of mixed ancestry, because less than 67% of their ancestry came from any single population. These STs were found at all locations but were less likely to be found in China where only 6% of STs had mixed ancestry [$\chi^2$ test, $P<0.05$]. The only significant difference between the Pacific STs
[Auckland, Samoan and Fijian STs grouped together] and the STs from other parts of the world were the CC75 outlier strains [coloured blue in Figure 5.38A]. STRUCTURE found that ST152, common in the Mali SA population, shared a portion [9%] of its ancestry with the CC75 outlier strains.

All of the remaining six sources of ancestry were present in each SA population, but the proportions of STs that belong to each ancestral population differ significantly in each country [Figure 5.38C]. For example, the Oxford SA population contains a large number of STs that belong to ancestral population four [coloured pink in Figure 5.38A, and 5.38B]. These STs correspond with STs found in CC30, a CC that is found globally. The Oxford SA population contains a large number of novel STs within CC30 [ST31, ST32, ST33, ST37 and ST38] that are not present in the other SA populations. Thus the Oxford SA population contains a significantly greater proportion of ancestral population four strains compared with SA populations from other countries [$\chi^2$ test, $P<0.05$].
Figure 5.38 The ancestral populations of the global SA population STs [Oxford, n=73; China, n=33; Mali, n=20; Auckland, n=62; and Pacific, Samoa and Fiji combined, n=43] derived by STRUCTURE; [A] the proportion of each ancestral population in each SA population; [B] the ancestry of each individual ST in the global SA population; [C] the proportion [%] of membership each ancestral population for STs from each country.

<table>
<thead>
<tr>
<th>Ancestral</th>
<th>England</th>
<th>China</th>
<th>Mali</th>
<th>Pacific(^1)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [red]</td>
<td>7</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>2 [blue]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4*</td>
<td>2</td>
</tr>
<tr>
<td>3 [orange]</td>
<td>8*</td>
<td>27</td>
<td>25</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>4 [pink]</td>
<td>24*</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>5 [brown]</td>
<td>14</td>
<td>33*</td>
<td>10</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>6 [green]</td>
<td>9</td>
<td>0*</td>
<td>10</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>7 [grey]</td>
<td>16</td>
<td>15</td>
<td>20</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>mixed ancestry</td>
<td>22</td>
<td>6*</td>
<td>20</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>

\(^1\)includes SA samples from Auckland, Fiji, Samoa and Tonga

\(^2\)comparison between countries for each ancestral population, X\(^2\) test, P<0.05

5.4.3.4 No difference between the SA population in Pacific countries and the SA population in Mali

The two dimensional multidimensional scaling [2D-MDS] plot of the global SA population [with outlier CC75 STs removed] shows considerable overlap between SA populations from different countries [Figure 5.39]\(^{149}\). ST152 present only in the SA population from Mali was distinct from all other populations.
the other STs. There were no significant differences between the SA populations from each country.

![Figure 5.39 The 2D-MDS plot of SA STs isolated in England [n=73], China [n=33], Mali [n=20] and the Pacific [n=87]. ST152, isolated in Mali was separated from the remaining STs.]

5.4.3.5 Estimating the Molecular clock rate of SA

The similarity of SA populations from geographically diverse countries in the absence of admixture through modern human migration suggests that dominant SA strains dispersed globally in the past. The phylogenetic tree contains information about the demographic history of the SA population [changes in the effective population size]. The molecular clock rate is used to estimate the time when changes in the effective population size of SA occurred.

The molecular clock rate is the number of nucleotide substitutions that affect the genome per year; but is not known for SA. Ochman and Wilson derived a phylogenetic tree based on the nucleotide sequences of 16sRNA from a number of different bacterial species and calibrated nodes of the tree using time scales estimated from the fossil record and knowledge of Earth’s geological history. For example, they estimated the evolution of aerobic bacteria to have occurred no later than 2,000 million years ago, when atmospheric oxygen levels are estimated to have risen; they estimated the evolution of rhizobial species of bacteria occurred at a time when the fossil record indicates that seed plants evolved. Using the calibration points they identified, it was possible to recreate a phylogenetic tree that included SA and one of its closest related species S. epidermidis, to determine when these two species diverged.

The 16sRNA sequences from reference strains of staphylococcal species and from the bacterial species used by Ochman and Wilson were downloaded from the ribosomal database project.
hosted at the University of Michigan [http://rdp.cme.msu.edu/]. These sequences were aligned using Geneious 3.5 and a phylogenetic tree was recreated using TREEPUZZLE 5.2 under the assumption of a molecular clock \(^{127}\). The assumption of a molecular clock results in a tree where the tips of the tree have branch lengths that are equidistant from the root. This assumption allows the time of divergence between two taxa to be estimated. The tree of the 16sRNA sequences is shown in Figure 5.40. The probability value of the tree reconstructed under the molecular clock \([\text{Ln} \ L] = -9950\) was less than the tree that was reconstructed at the same time without the assumption of the molecular clock \([\text{Ln} \ L] = -9859\).

The time of divergence between two taxa was estimated by determining the age of the ancestral node at the root between the two taxa. The age of the ancestral node for escherichia and salmonella, estimated to have diverged 140 million years ago, was 0.0192. Thus, SA diverged from \textit{S. epidermidis} between 5 and 10 million years ago, [the node age of SA and \textit{S. epidermidis} was 0.0012] and staphylococci diverged from Bacillus spp. and Listeria spp. approximately 300 million years ago.
Figure 5.40 The phylogenetic tree of staphylococci and other bacteria used by Ochman and Wilson to date the time of divergence of bacterial species. The tree was reconstructed from 16sRNA nucleotide sequences using TREEPUZZLE 5.2 under the assumption of a molecular clock. The age of the ancestral node for escherichia and salmonella [blue] was used to determine the divergence of SA from *S. epidermidis* [orange].
The molecular clock rate for SA was determined from the time taken for SA and *S. epidermidis* to diverge from their common ancestor. The entire genomic sequences for *S. epidermidis* ATCC12228 and SA strains MW2, RF122, and COL and Newman [http://cmr.jcvi.org/cgi-bin/CMR/CmrHomePage.cgi](http://cmr.jcvi.org/cgi-bin/CMR/CmrHomePage.cgi) were aligned using MAUVE, the average pairwise distance between *S. epidermidis* and the SA genomes was 41% [Figure 5.41](#)\(^1\). Thus SA has accumulated changes to 20.5% of its genome over 7.5 million years, which provided a molecular clock rate of \(2.73 \times 10^{-8}\) year.

![Figure 5.41 The alignment of genomes of SA RF122 (above) and *S. epidermidis* ATCC12228 (below) using MAUVE. The pairwise distance between these genomes was 42.2%.

5.4.3.6 The SA population is young and the effective population size of SA has increased recently

TREEPUZZLE was used to reconstruct the phylogeny of the global SA sample under the assumption of a molecular clock [Figure 5.42A](#)\(^2\). The assumption of a molecular clock creates a tree which has branch lengths that are proportional to the time since divergence from a common ancestor. The molecular clock tree had a lower likelihood score than the tree that was reconstructed without this assumption, which allows for different rates of evolution in different parts of the phylogeny. The recent burst of evolution of SA, indicated by the short branch lengths for many of the strains, reflects an increase in the effective population size of SA.

The generalised Bayesian skyline plot, showing the demographic history of SA [Figure 5.42B](#) was created using GENIE 3.0 from the molecular clock tree\(^3\). The molecular clock rate of \(2.73 \times 10^{-8}\) per year was used to date the changes in the population size in the past. A number of models of population change were tested; the model of stepwise expansion gained the highest likelihood. The stepwise model allows for a single change in a constant population size, which occurred approximately 3,500 years ago [Figure 5.42B](#), leading to a 10 fold increase in the effective population size.
Figure 5.42 The demographic history of SA: [A] The phylogeny of a global sample of SA created by maximum likelihood under the assumption of a molecular clock using TREEPZUZLE 5.2. The deep evolutionary branch to ST152 is shown in red; [B] Generalised skyline plot of effective population size relative to the present versus time in the past.
5.4.3.7 The origins of CC75 outlier strains and time of divergence

The existence of the divergent CC75 strains is intriguing; it is puzzling that the genetic diversity that they display is not seen in the much more homogenous conventional SA population. The phylogeny shown in Figure 5.13 indicates the age of the divergence of the CC75 outlier strains from the main SA population.

The nucleotide sequences of gene fragments from the open reading frame of coa from CC75 strains and conventional SA were compared to demonstrate that CC75 strains do not originate from a coagulase-negative staphylococcus species that has undergone recombination with SA. If the coagulase gene, coa, has not recently been acquired by CC75 strains then CC75 coa gene sequences should differ from those of conventional SA. The coa gene sequences were aligned using GENEIOUS 4.5, and regions of the alignment containing insertions and deletions were removed by editing the amino acid sequence. The maximum likelihood phylogenetic tree, reconstructed by TREEPUZZLE, indicates that the divergence shown by the CC75 MLST housekeeping genes is also displayed in the coa gene, confirming that CC75 strains diverged from conventional SA a relatively long time ago [Figure 5.43].

Figure 5.43 Coagulase gene [coa] phylogeny [midpoint rooted] indicating divergence between CC75 SA strains and conventional SA strains [SA to SAe].

The timing of the divergence of CC75 strains from conventional SA was estimated from phylogenetic trees reconstructed from nucleotide sequences from aroE and pta gene fragments. These phylogenies [Figures 5.44, 5.45] were reconstructed from nucleotide sequences of CC75 strains, predominant STs of conventional SA, and other staphylococcal species [S. capitis SK14, S. epidermidis RP62A, S. epidermidis ATCC12228, S. haemolyticus JCSC1435, S. hominis]
SK119, *S. saprophyticus* ATCC15395, *S. warneri* L37603; obtained from [http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi](http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi). The *aroE* gene fragments [261/456, 57% polymorphic sites] contain more diversity than the *pta* gene fragments [216/474, 46% polymorphic sites]; and both phylogenies show the close ancestral relationship between CC75 strains and conventional SA compared with other staphylococcal species. The same trees were reconstructed under the assumption of a strict molecular clock. The likelihood of the molecular clock tree reconstructed from *aroE* gene fragments did not differ significantly from the likelihood of the tree reconstructed without the assumption of a molecular clock; however, the likelihood of the molecular clock tree reconstructed from *pta* gene fragments was significantly less than the likelihood of the tree reconstructed without the assumption of a molecular clock [likelihood ratio test 30.99, 17 d.f., $P=0.02$].

The trees reconstructed under the assumption of a molecular clock consistently indicate that the hypothetical common ancestor to the CC75 strains and conventional SA strains is approximately 2.5 times younger than the common ancestor of SA and *S. epidermidis*. Thus, CC75 strains diverged from conventional SA strains more than one million years ago.
Figure 5.44 [A] Phylogeny of SA, CC75 strains and other staphylococcal species reconstructed from \textit{aroE} gene fragments; [B] the same tree reconstructed under the assumption of a molecular clock.
Figure 5.44 [A] Phylogeny of SA, CC75 strains and other staphylococcal species reconstructed from *pta* gene fragments; [B] the same tree reconstructed under the assumption of a molecular clock.
5.5 Discussion

The incidence of *S. aureus* bacteraemia [SAB] in New Zealand between 1996 and 1997 was the highest reported anywhere in the world. In North America the prevalence of invasive SA disease has increased exponentially since 2000, in parallel with the increase in prevalence of infection caused by methicillin-resistant SA strains including the USA300, ST8, SA strain. There is no single hyper-virulent clone of SA responsible for the high incidence of SAB in New Zealand; a wide range of SA strains with diverse ancestry cause infection in Auckland.

Examination of the population structure of SA in Auckland and in the Pacific region in relation to the global SA population is fundamental to the understanding of SA evolution which is central to the development of strategies to prevent illness caused by SA.

5.5.1 The SA population in Auckland

Analysis of the Auckland SA population, represented by the strains of SA obtained from 150 cases of SAB [blood population] and 94 nasal carriers [nasal population], identified a diverse range of STs. A genetic outgroup of four SA strains was identified within the nasal SA population, and these similar strains are closely related to strains which belong to CC75, previously identified in Australia and Cambodia. SA strains do not have equal propensity to cause infection in Auckland; 62 unique STs were identified among the 244 strains, but just 5 STs caused half of all infections.

5.5.1.1 Origin of novel STs in Auckland

34/244 [14%] of the isolates in Auckland are novel, they have not previously been described. This finding is consistent with previous studies; Fan et al, found that 15% of strains isolated from Chinese children with nasal colonisation were novel. In Auckland, novel STs are dispersed throughout the phylogeny and no genetic background is more likely to give rise to novel STs. The predominant CCs in Auckland are similar to the predominant CCs in China [CC1, CC5, CC15, CC30, CC59, CC88 and CC121 are common in both countries]. In Auckland and China novel STs arise within predominant CCs; within any given CC the novel Auckland ST[s] differ from the novel ST[s] in China. Thus the novel STs in Auckland have evolved in Auckland.

Recombination is an important driver of the evolution of SA; over half of the novel STs arose following a recombination event. Overall, the influence of recombination on any nucleotide substitution is twice that of mutation. The ratio of recombination to mutation at each nucleotide did not differ between the Auckland SA population and simulated populations of SA obtained from the STs entered into the MLST database.
Population genetic methods, used in the current study to determine the relative influence that recombination and mutation \([r/m]\) have on diversification at each nucleotide, confirmed the low rate of recombination in SA populations suggested by Feil et al. Comparison was made between simulated samples of SA and *N. meningitidis*, *S. pneumoniae* [organisms with high rates of recombination]\(^{159,160}\). *N. meningitidis* had rates of both recombination and mutation that were almost 10 times greater than SA; *S. pneumoniae* had rates of recombination that were approximately 20 times greater than SA.

### 5.5.1.2 No difference between blood and nasal SA populations in Auckland

There was no difference between the STs of SA that caused invasive disease and those that caused nasal colonisation among people living in Oxford\(^ {112}\). In the current study a number of methods confirmed that SA strains that cause invasive SA disease are not a more virulent subset of strains that circulate asymptptomatically in the population. In Auckland, the significantly higher proportion of ST1 strains that caused SAB [17\%] compared with nasal colonisation [4\%] was the only significant difference between the SA population that caused SAB and the SA population that caused nasal colonisation. Visual comparison of the SA populations, performed by multi-dimensional scaling, showed no difference between blood and nasal SA populations. LibShuff, ANOSIM and PERMANOVA failed to identify any significant difference between the two populations. Finally, the ancestral sources, defined by STRUCTURE, of each population did not differ, with the exception of the CC75 outlier strains present only in the nasal carriage sample, whose source of ancestry was entirely different from the remainder of the strains.

The analysis performed in this thesis was based on a representation of the core genome of SA; the additional analysis of *ssl7* gene sequences did not enable further distinction between SA strains. It has been shown that some virulence genes are not distributed evenly between SA that cause invasive disease and nasal colonisation\(^ {113}\). Furthermore, substantial differences were found by comparison of a genome of SA that caused disease with a genome of SA from the same ST that caused nasal colonisation\(^ {95}\); although it is unclear whether the differences between the two genomes involved the determinants of virulence. The SA strains collected for this thesis provide an opportunity to compare the genomes of multiple SA strains, of the same ST, that have caused SAB and nasal colonisation. Such a study would expand upon the findings of Peacock et al and Sivaraman et al, and provides an opportunity to comprehensively examine the determinants of invasive disease.

### 5.5.1.3 Difference in SA STs in different ethnic groups

Infection by SA is more common in Māori and Pacific people than people of other ethnicities as is infection with Western Samoan phage pattern methicillin-resistant SA [WSPP MRSA] [Chapter 2]\(^ {36,37}\). In order to investigate whether ethnic variation in the prevalence of SAB is caused by a predisposition to infection with specific virulent SA STs, such as WSPP MRSA, the SA population that caused infection in Māori and Pacific people was compared with the SA populations that
caused infection in people of other ethnicities. The sequence type of WSPP MRSA is invariably ST30, and infection with ST30 strains or closely related strains in CC30, was not more common in Māori and Pacific people. Indeed, the SA population obtained from Māori and Pacific people was not significantly different to the SA population obtained from people of other ethnicities. However, CC121 strains were only isolated from Māori or Pacific people. ST121 strains are common causes of nasal colonisation in Asian countries; and it is possible that a methicillin-susceptible CC121 strain that causes skin and soft tissue infection in Germany, is globally disseminated. However, the finding of a novel ST121 variant [ST11131] in Auckland, Tonga, Fiji and Samoa suggests that CC121 strains have adapted to people in the Pacific region. ST11131 differs from ST121 at the gmk locus, following a recombination event. The new gmk allele was not identified in the remainder of the Auckland SA population and might have its origins elsewhere.

5.5.2 Comparison of the Auckland SA population with SA populations in Fiji, Samoa and Tonga

Samples of SA were obtained from people with wound infection in Samoa and from the public hospital microbiology laboratories in Fiji and Tonga. Each sample provided an excellent representation of the total SA populations in each country which were compared with the SA population in Auckland. Despite their small size, the evolution of novel STs is common in Samoa and Fiji; in Samoa 9% of STs are novel and in Fiji 16% of STs are novel.

5.5.2.1 Difference between the frequencies of predominant STs in Pacific countries

The frequency of the predominant STs found in different locations varies, although there is consistency between countries in the same region. For example, the frequency of the predominant CCs in France, England and Algeria [SA sample collected along the Mediterranean coast] are similar but differ from the frequency of predominant CCs in China and Cambodia, which are also similar [Figure 5.43]. There is no such regional consistency in the Pacific SA population; comparison of the populations of SA from Auckland, Samoa and Fiji demonstrate striking differences in the frequency of predominant STs between countries.
Figure 5.43 The proportion of common CCs in global populations of SA.

ST1 is a common cause of infection in Auckland and Samoa but is not found in Fiji. This finding is surprising, ST1 strains caused over 10% of the infections in Samoa and Auckland; although CC1 strains are less common in other parts of the world [Figure 5.43]. The absence of ST1 in the Fijian sample is not likely to be due deficiencies in the sample, which included SA that had caused disease, predominantly skin and soft tissue infection. Firstly, ST1 strains are a common cause of skin and soft tissue infection in Samoa and Auckland. In Auckland, 10/24 episodes of SAB caused by ST1 strains were associated with skin and soft tissue infection. In Samoa all of the SA isolates were associated with skin and soft tissue infection. Secondly, the sample of SA obtained from Fiji was an excellent representation of the total sample by rarefaction analysis; and the sample would have identified any ST with a frequency of more than 5%. Finally, the Fijian SA population contains STs that belong to CC1, indicating that this genetic background is uncommon in Fiji, but not absent.

ST30 strains preferentially cause skin and soft tissue infection in Fiji and ST1 preferentially cause skin and soft tissue infection in Auckland and Samoa. The reasons for variation in the prevalence of infections caused by different STs in neighbouring, small Pacific countries is likely to be complex. Environmental factors are not likely to play a large role, Fiji and Samoa are geographically close and their climates are similar. Likewise, human factors are unlikely to be important despite differences between the ethnicity of each country’s population. Almost all indigenous Fijian people are of Melanesian descent whereas Samoan people are of Polynesian descent; however, the population of Auckland is diverse and analysis of the entire Pacific region SA population found no ethnic predisposition to infection with ST1 or ST30 strains. Furthermore, some mixture between Melanesian and Polynesian people occurred as Polynesian people
migrated into the Pacific. The high proportion of the Fijian population that are Indian also makes it unlikely that absence of ST1 strains in the Fijian SA population is due to reduced adaptation to a Melanesian host.

5.5.2.2 The SA populations in neighbouring Pacific Island nations are similar

Despite variation in the frequency of predominant STs in each SA population, the degree of genetic diversity of each SA population is very similar, and almost all SA genotypes are present in each population. Significant differences between SA populations were detected only when the sample size was increased by combining SA populations: the sample of SA obtained in Auckland is significantly different to the sample of SA obtained in all three Pacific Island nations; and the sample of SA obtained from Samoa is significantly different to the sample of SA obtained from other Pacific nations [including New Zealand]. There are a number of indications that the differences detected are small. Firstly, Simpson’s index of diversity did not differ between populations. Secondly, the ANOSIM R values were close to zero; the SA genotypes in Samoa [or Auckland] were only marginally more similar than comparison with the SA genotypes found in the other countries. Thirdly, there was only a non-significant trend towards a difference between the genotypes found in Fiji and the other Pacific countries. Finally, visual representation of the SA samples from each Pacific country by multi-dimensional scaling showed that the SA genotypes within each population overlapped, with the exception of the occasional outlier genotype in Auckland and Samoa.

5.5.2.3 CC121 strains are common in the Pacific

In Auckland, CC121 strains were only identified in Māori and Pacific people. CC121 strains were more common in SA populations in the Pacific Island nations than in Auckland. This finding raises the possibility that these strains have recently entered the Pacific, but have not yet become widely disseminated in the population of Auckland; or that these strains have adapted to cause infection in hosts of different ethnic groups. Further testing is required to determine whether there is a single clone of ST121 that has adapted to Māori and Pacific people. However, the novel strain ST11131, which is a member of CC121, caused infection in Auckland, Samoa, Fiji and Tonga and represents a successful novel ST of CC121 possibly adapted to Māori and Pacific people.

5.5.2.4 Evolutionary relationships derived from point mutation identified divergent SA strains in Auckland and Fiji

The phylogeny of the Pacific SA population [Figure 5.29] shows that a hypothetical ancestor of diverged form the main population of SA and a small number of divergent outlier SA strains found in Auckland and Fiji. These strains were compared to the MLST database and are similar to CC75 SA strains. In the MLST database, some CC75 strains contain a conventional $aroE$ allele among six divergent loci. This anomaly is likely to be the result of a mismatch between the PCR
primers\textsuperscript{170}; this was not encountered in the current study as new PCR primers were used to amplify and sequence the \textit{aroE} nucleotide sequence. Although the CC75 strains differed greatly from the main population of SA found in the Pacific; they occupy the same niche as conventional SA [anterior nares in Auckland, wound infection in Fiji], they are coagulase positive and produce staphyloccocal DNase. Furthermore, despite the high degree of divergence in their housekeeping genes, the CC75 outlier strains’ 16sRNA gene sequences have not diverged from those of conventional SA.

CC75 strains were also found colonising wounds in remote Aboriginal populations and in a single nasal carrier in Cambodia\textsuperscript{120, 170}. Ng et al, postulated that CC75 SA may have adapted to tropical regions but their existence in Auckland indicates that this is unlikely\textsuperscript{170}. It is now evident that these strains are widespread, but are likely to be of low virulence; previous studies that have focussed on invasive disease isolates have not identified CC75 strains causing invasive disease. The evolutionary history does not indicate that CC75 strains are the ancestors to conventional SA strains, but they have diverged markedly from the main SA population; they are likely to be living fossils. The phylogenies of staphyloccocal species indicate that the common ancestor for CC75 strains and conventional SA strains is younger than the common ancestor for SA and \textit{S. epidermidis} [Figures 5.44, 5.45]. Thus CC75 strains diverged from the main SA population well over one million years ago. It is surprising that the main SA populations of Auckland, Samoa and Fiji are so genetically similar in relation to the diversity shown by the CC75 strains.

5.5.3 The global SA population - novel STs are common

Evolution of novel STs is common in small Pacific nations, but the SA populations are very similar and lack substantial geographical structure. In order to determine whether this finding was unusual, comparison was made with SA populations in Oxford, China and Mali. Any detectable difference between SA populations in Pacific countries is likely to represent the differences that arise from the evolution of novel STs from the predominant SA genotypes within each country. The evolutionary relationship between SA populations from Oxford, Mali, China and the Pacific indicates that novel STs have evolved from the dominant STs in each country. For example, the Oxford SA population contains numerous SA genotypes not present in other countries. ST30 was the most frequent ST in the Oxford sample and novel variants of ST30 were particularly common. Thus, each SA population contains some geographical structuring, evident in the novel STs that have evolved in each country. Harris et al also demonstrated geographical structure in ST239 MRSA strains by comparing polymorphism across the genomes of 63 ST239 strains obtained from Europe, Asia and South America\textsuperscript{188}. The phylogeny of ST239 MRSA strains reconstructed from polymorphisms in the core genome demonstrated relationships between taxa on the basis of geography and time of isolation.

The novel STs found in each country are dispersed throughout the phylogeny, and it does not appear that any dominant genotype gives rise to novel ST more often than any other dominant
genotype. The ability to identify novel STs is dependant on the frequency of the predominant ancestral genotype in the population. ST30 and novel variants of ST30 are common in Oxford and the Pacific; in contrast, ST30 is rare in China, and novel ST30 variants were not identified in the sample of SA from China. The rarefaction curves for Auckland and the Pacific indicate that the SA sample size required to identify all novel STs in each country is huge, and a much larger sample from China would be expected to identify novel ST30 variants.

5.5.3.1 Geographically diverse SA populations are similar

STRUCTURE verified that there is no substantial geographical clustering of SA populations, despite evolution of unique novel STs in each location. The proportions of ancestry in each SA population varied; however, with the exception of CC75 outlier strains which form a distinct ancestral population, each ancestral source is represented in all modern SA populations. There are significant differences in the proportions of each ancestral source between SA populations and this indication of minor geographical structure is caused by variation in the frequency of predominant STs and the evolution of novel STs in each SA population.

Despite inability to identify all novel STs in each SA population, the sample size of SA from each country is more than adequate to allow comparison of the population structure of SA between each country. Predominant strains [ST1, ST5, ST8, ST15, ST30 etc.] were present in all locations. However, it is remarkable that, with the exception of ST152, which is common in Mali but not found in other SA populations, the genetic diversity of SA populations from geographically diverse places is the same. Ruimy et al concluded that the similarity between the SA populations in Algeria and France is due to frequent modern human migration between these two nations and that the difference between SA populations in France and Cambodia is due to infrequent human migration between these two nations. Nubel et al did not find evidence of recent intercontinental spread of ST5 SA; however, Harris et al, found evidence of intercontinental transmission of ST239 SA.

5.5.4 A model of SA evolution and dispersal

The analysis of geographically diverse SA populations raises several anomalies. Firstly, the contemporary SA population is relatively young and has undergone a recent population expansion [Section 5.4.3.6]. Secondly, contemporary STs are highly related compared with divergent CC75 STs [Section 5.4.3.7]. Thirdly, geographically diverse SA populations are very similar and, although the frequency of predominant STs varies between countries, lack substantial geographical structure [Sections 5.4.3.3, 5.4.3.4]. A plausible model of the evolution of SA, developed from the results presented in this Chapter is shown in Figure 5.44. The model suggests that the evolution of SA occurred in several phases:

Ancestral phase: SA diverged from other staphylococci several million years ago.
**Isolation phase:** SA dispersed with early human migration to distinct, relatively isolated environments. Diversification occurred within each isolated SA population.

**Bottleneck phase:** A genetic bottleneck occurred, probably as a result of a global selective sweep that purged diversity from the global SA population. After the genetic bottleneck SA diversified to form several related STs and clonal complexes.

**Rapid migration phase:** The contemporary SA population has become mixed by modern human travel and migration. Selection pressure [e.g. antimicrobial usage] is a critical component of SA evolution and selective sweeps are expected to produce SA strains which disperse globally.

**5.5.4.1 The ancestral phase of SA evolution**

SA has co-evolved with humans. Ng et al used evolutionary relationships derived from conserved staphylococcal genes to demonstrate that SA originated at approximately the same time as *S. epidermidis* and *S. capitis* [170]. In this Chapter, 16sRNA nucleotide sequences were used to estimate that SA diverged from *S. epidermidis* between 5 and 10 million years ago; at the same time [Miocene epoch] that early hominids began to appear.

Since those origins SA has developed an array of immune evasion factors that have striking adaptations to the human immune system. For example, staphylococcal superantigen-like protein 7 [SSL7] has affinity for human and primate IgA but does not bind to IgA from other animals such as cows, sheep and mice [175]. The *ssl7* gene is found in all SA strains and *ssl7* gene polymorphisms mirror consistent changes in MLST housekeeping genes such that the *ssl7* gene sequence could be used to accurately predict the CC of an SA isolate. The specificity of immune evasion factors such as SSL7 for the human immune system suggests that SA has been a long-standing human pathogen or commensal.

**5.5.4.2 The isolation phase of SA evolution**

The evidence that SA is predominantly a human pathogen that has had a long-standing relationship with humans indicates that SA dispersed around the globe with early humans. A large number of relatively isolated SA populations had ample time to develop marked diversity. Evidence of this diversity is found in the CC75 strains and ST152, which represent SA strains that diverged from contemporary SA strains during the ancestral phase or isolation phase [pink wedge in Figure 5.44]. These STs are not prominent on a global scale; ST152 is a common cause of nasal colonisation in Mali but infrequently causes infection elsewhere [119]; CC75 strains do not appear to be a common cause of infection outside of Aboriginal communities in Northern Australia [170]. It is likely that other older SA STs persist, but are infrequent causes of SA disease and await discovery.
Figure 5.44 Evolutionary model of SA. SA dispersed globally with early human migration and diversified in isolation from other SA populations, remnants of this diversification persist today [pink wedge; e.g. CC75 strains]. A genetic bottleneck occurred prior to recent population expansion and contemporary SA populations have become mixed.
Genetic drift rapidly affects SA populations. The rate of change in the total SA genome in relation to the rate of change in the MLST loci is not known, but the number is likely to be large. Analysis of the single nucleotide polymorphisms of 63 ST239 strains collected over only a few decades identified 4310 variable sites in the core genome. ST239 SA is under selection pressure from antimicrobial use that was not present in ancient SA populations, yet the majority of the polymorphisms arose due to random mutation and genetic drift. If the rapid molecular clock rate of ST239 MRSA and ST225 MRSA is indicative of the molecular clock rate of other SA, then it is impossible that isolated SA populations in diverse locations did not develop diversity and striking geographical population structure. The expected degree of diversity, suggested by the divergent CC75 strains, is not present in contemporary SA populations and has been lost.

5.5.4.3 The bottleneck phase of SA evolution

The effective population size of SA has increased dramatically recently indicating that the SA population has gained diversity rapidly. The relative homogeneity of the present SA population indicates that prior to the increase in population size there was little diversity in the SA population. Thus, a genetic bottleneck has occurred recently in the evolutionary history of SA as a result of a selective sweep of an SA strain[s] with an advantage over other SA strains that purged diversity from SA populations. The selective sweep occurred on a global scale and can only have occurred in the presence of human movement between previously isolated populations. The first occasion when isolated human populations began to move rapidly and in large volumes was approximately 300 years ago when ocean travel and human colonisation began. For a short time after the genetic bottleneck, the main SA population consisted of relatively few closely related SA STs, which have given rise to contemporary SA clonal complexes. It is unlikely that a global selective sweep would purge all divergent STs, and the identification of divergent STs is not in conflict with the model.

The skyline plot of the global SA population clearly demonstrates the recent increase in effective population size, but the time scale is not consistent with the model. It is important to note that there is potential for error in the estimation of the time frame of the population expansion. Firstly, the molecular clock rate [nucleotide mutations / year] of SA was estimated from the divergence of SA from S. epidermidis, which was estimated to have occurred 7.5 million years ago; but this estimate falls within a much wider range. The divergence time of SA from S. epidermidis could be long as 20 million years or as short as 1 million years. The molecular clock rate derived for SA [2.73 x10^{-8}] is similar to the estimated molecular clock rate of Escherichia coli. Secondly, the skyline plot was derived from a maximum likelihood tree reconstructed under the assumption of a molecular clock [consistent rate of mutation along all branches], which was not as likely as a tree which allowed variable rates of mutation along different branches in the tree. Finally the model of population expansion for the global SA population allowed for a single step-wise change in a constant population size, and natural populations seldom remain at a constant size from one generation to the next.
Recently, Nubel et al. estimated the molecular clock rate of ST225 MRSA from analysis of genome fragments from strains collected between 1994 and 2007. Their estimated rate [2 x 10^6/nucleotide/year] is several orders of magnitude greater than the estimate derived from the divergence of SA from S. epidermidis; but is consistent with another recent estimate. This molecular clock rate dated the recent increase in effective population size to the past 10 years, which is inaccurate but more consistent with the proposed evolutionary model. If the SA population expansion has occurred more recently, for example within the past 100 years, then the genetic bottleneck might have been caused by the widespread use of penicillin. Resistance to penicillin may have developed in a small number of related SA strains, which then replaced more divergent penicillin susceptible strains.

5.5.4.4 The rapid migration phase of SA evolution

Recent mixture is the only simple explanation for the finding that geographically diverse SA populations lack substantial differences. Since the bottleneck phase of SA evolution, populations of SA have become increasingly admixed and have failed to develop geographical population structure, despite an increase in diversity. The current variation in frequency of predominant SA STs in different countries might represent residual geographic population structure that originated after the genetic bottleneck. Thus, differences between the frequency of ST1 and ST30 infections in Samoa and Fiji could be explained by founder effect: ST1 has been more common in Samoa and ST30 has been more common in Fiji, since pre-existing diversity was purged by a global selective sweep that occurred after European colonists migrated to Pacific Island nations.

Most of the increase in SA population diversity involves the evolution of single locus variant novel STs from predominant STs. The finding of frequent evolution within SA populations indicates that SA populations are not made up of stable predominant STs. Most of the single locus variant strains identified in Oxford 10 years ago were novel STs and were isolated only once. The absence of subsequent entries of the same STs in the MLST database and their absence in other SA populations suggest that novel STs do not often persist. Similarly, most of the novel variants identified in the Pacific will fail to persist. However, as expected, some novel STs do persist; in the Pacific SA population ST11131 is a successful novel variant of ST121 and causes infection in different people in different Pacific countries.

The incidence of infections caused by MRSA has increased in many parts of the world, and infections caused by methicillin-susceptible SA are probably also increasing but less frequently reported. It is likely that modern selection pressures, recombination and rapid migration have played an important part in the increased prevalence of SA disease. The recent evolution of a global pandemic SA population indicates that selective sweeps now operate on a global level. The emergence of the USA300 MRSA strain in North America followed by relentless international spread represents a selective sweep as this strain gains a foothold in other SA populations. It is likely that selective pressure on SA populations is higher than ever before; the successful USA300 and ST239 MRSA strains have arisen, partly as a consequence...
of antimicrobial pressure\textsuperscript{184, 188}. Increased human population density and air travel allow frequent opportunity for novel SA strains to evolve and disperse in new environments.

The rapid global spread of novel SA strains also provides some evidence of the high transmissibility of SA, which is a fundamental requirement of this model. It is likely that SA strains spread rapidly through human populations; a very high proportion of infants become transiently infected during the first year of life\textsuperscript{82, 101}; and asymptomatic colonisation with different SA strains was recently demonstrated to occur over a short time period\textsuperscript{84}.

5.5.4.5 Alternative models of SA evolution

A number of alternative models were explored during the development of the evolutionary model, but were quickly discounted. For example, a model that explains the similarity of geographically diverse SA populations by the migration of closely related STs with early human migrants is not plausible because evolution in isolation would give rise to variation in predominant STs between countries and older SA populations in Africa would be considerably more diverse than younger populations in the Pacific. Furthermore, genetic drift occurring in isolation for tens of thousands of years would have increased the diversity of the modern SA population immensely.

One alternative model was more difficult to discount. The model makes the assumption that SA has recently become a human pathogen from an animal reservoir or from the environment. This provides an explanation for the lack of diversity in the human-associated SA population. As an example, the diversity of influenza A viruses that infect humans is minimal in comparison with the diversity of influenza A viruses that infect birds. If a SA ST recently adapted to cause human infection it is expected that the diversity of human SA would be minimal but increasing, as is the case. The diversity between CC75 SA strains and the main SA population could have arisen if SA adapted to different animal hosts over one million years ago and then diversified. Thus, this model requires that SA had crossed species and adapted to humans on at least two occasions. Furthermore, the marked adaptation of SA for humans suggests that SA infection is not zoonotic.

The distinct SA strains that have adapted to birds, cows and other animals are not markedly divergent from human SA strains\textsuperscript{197, 198}. The description of adaptation of human SA to poultry demonstrates the efficiency with which SA adapts to a new host and it is probable that distinct strains of SA found in other animals have arisen by similar mechanisms\textsuperscript{198}. Strains of SA that cause infection in animals have probably originated from humans; although, Plomett and Wilson found evidence that SA infects wild animals\textsuperscript{199}. They found that SA isolated from wild animals [such as beavers, foxes, and raccoons] were of different serotypes when compared with SA isolated from humans and domestic animals. Furthermore, inability to phage type the wild animal SA strains also suggests they differ substantially from the other SA strains. Thus, it is possible that Plomett and Wilson found evidence of the marked diversity in animal reservoirs of SA that exists if SA has recently become a human pathogen; however, more recent studies indicate that
the diversity of animal-adapted SA populations mirrors the restricted diversity of human SA populations.

5.6 Summary

The population genetic study of the SA populations in New Zealand and some Pacific Island nations has found evidence of frequent evolution of SA and has identified a lineage of SA strains [CC121], including a successful novel ST, that have adapted to cause infection in Māori and Pacific people. An evolutionary model was derived to explain several intriguing findings, including the existence of divergent SA strains among an otherwise homogenous SA population and the striking similarity of SA populations from diverse countries. It is important that future research investigates the selection pressures that shape modern SA populations, the increasing diversity of SA will result in frequent evolution of virulent strains and their rapid global dispersal will pose an increasing threat to human and animal health.
Part C: Human immune responses to staphylococcal toxic shock syndrome toxin, and staphylococcal superantigen-like proteins 3, 7, 9 and 11.

Chapter 6: Demographic variation in the immune responses to staphylococcal toxic shock syndrome toxin and lack of neutralisation of specific anti-SSL7 antibody

6.1 Introduction

Part C aims to examine the influence of the host immune response on the development of SA disease; in particular to answer whether ethnic variation in the prevalence of SAB might be due to variation in host immune response. There are currently no standard methods to quantify a person’s immunity to SA.

There are a large number of virulence factors that are potentially important in the development of SA infection and maintenance of disease, but no single virulence factor has consistently been shown to be associated with invasive disease, with the exception of staphylococcal toxic shock syndrome [TSS], which is only caused by *tst* positive strains of SA strains. Some evidence suggests that TSST-1 is important in SA bacteremia [SAB] and anti-TSST-1 antibody levels will be analysed in Part C. The staphylococcal superantigen-like proteins [SSL] are putative virulence factors in staphylococcal disease and several SSL proteins appear to have important roles in immune evasion. Antibody levels against these staphylococcal toxins might, therefore, provide further insight into host susceptibility to SA disease. Serum antibody levels to SSL proteins [SSL7, SSL3, SSL9 and SSL11] were quantified in patients with invasive staphylococcal disease.
6.1.1 Are effective vaccines against SA available?

There are no vaccines against SA currently in widespread use, yet the demand for an effective vaccine is high. Serious disease caused by SA is common, antimicrobial resistance is ever increasing, and the morbidity and mortality rates remain high, despite appropriate treatment. Vaccines against SA were used to treat skin infections and to reduce peri-operative infection in the era prior to antibiotic development. However, no vaccines have been supported by published randomised-controlled clinical trials. Only two vaccines have recently been studied in humans.

6.1.2 Vaccination against staphylococcal alphatoxin did not reduce SA infection in high-risk patients

Staphypan Berna vaccine consists of undefined extracts from whole killed strains of SA and S. epidermidis, together with staphylococcal alphatoxin, a pore-forming protein that is expressed during post-exponential growth and is able to lyse red blood cells, macrophages and lymphocytes.

In 1991, Poole-Warren et al reported the results of a 2 year controlled trial of the Staphypan Berna vaccine in 124 stable patients receiving chronic ambulatory peritoneal dialysis. The study was conducted in Australia and New Zealand; and patients were randomised to receive six injections of vaccine a week apart, followed by four booster injections every 12 weeks, or saline placebo. The vaccine did stimulate an immune response, serum anti-staphylolysin antibodies [a crude measure of anti-alphatoxin antibodies] rose in vaccinated patients and serum anti-SA antibody levels also increased. However, vaccination offered no protection against peritonitis, intra-peritoneal catheter infection or nasal carriage of SA. The vaccine had no impact on the growth rate of SA in dialysate that had been drained from the peritoneal cavity of vaccine or placebo recipients.

In 2002, Zachrisson et al reported a study that found a positive response to vaccination with Staphypan Berna in a small number of patients with chronic fatigue syndrome and fibromyalgia. The authors determined that the vaccine contained high levels of lipase, alphatoxin, enterotoxin B, toxic shock syndrome toxin 1 [TSST-1] and cell wall antigens. There was a significant improvement in the symptom scores of 14 vaccine recipients when compared to 14 participants who received placebo.

A positive response to vaccine correlated with neutralisation of alphatoxin function by serum collected prior to vaccination. This result suggested that participants who were able to produce effective immune responses to alphatoxin derived benefit from immunologic boosting. It is not known whether the response was related to an anti-SA effect or a non-specific response to vaccination. Furthermore, the results have not yet been confirmed in larger studies.
6.1.3 Vaccination against the SA polysaccharide capsule did not reduce SAB

Most clinical isolates of SA have a polysaccharide capsule, and these are usually of three common serotypes: 5, 8 and 336. The capsule resists opsonisation and SA strains that had serotype 5 and 8 capsules were more virulent in a murine infection model. In 2002, Shinefield et al. reported the results of a successful trial of a vaccine containing type 5 and type 8 capsular polysaccharide [StaphVAX]. In this trial, vaccine polysaccharides were conjugated to a non-toxic variant of pseudomonas exoprotein A. 894 people requiring haemodialysis were randomised to receive the vaccine and another 910 received a placebo injection. The results of this study showed partial protection against SAB that was not maintained beyond 40 weeks. A second study was unable to confirm these findings. There are several reasons why a vaccine against polysaccharide capsule might not have provided a sustained high level of immunity. Different capsule serotypes exist and strains that have lost their capsules remain virulent. A recent study demonstrated that vaccination with StaphVAX had no influence on the incidence of nasal colonisation.

6.1.4 Do immune responses to infection protect against SA infection?

The inability of vaccines to prevent SA disease is not surprising when one considers the enormous array of virulence factors that SA strains possess. In broad terms, these virulence factors fall into three categories: factors that assist with adherence to host cells; factors involved in immune evasion; and factors that aid in tissue destruction and spread. There is mounting evidence that immunity to these virulence factors is important in susceptibility to SA disease.

6.1.5 People who are immune-suppressed are at increased risk of SA infection

Case-control studies have shown that nasal colonisation with SA is more common in people with comorbid illnesses that result in impairment of the immune system. Specifically the illnesses that have been studied are insulin dependent diabetes, HIV infection, intravenous drug users and people with end-stage renal failure requiring dialysis. These illnesses, amongst others, are also classical risk factors for the development of invasive disease caused by SA.

6.1.6 Differences between persistent non-carriers and persistent carriers of SA

Longitudinal studies of healthy volunteers have found that approximately 30% of people do not develop nasal colonisation with SA and have been termed persistent non-carriers. These studies defined three types of healthy adults: persistent non-carriers, intermittent carriers and persistent carriers. In a study of experimental nasal colonisation, nasal inoculation with one of four strains of SA was less likely to persist in non-carriers when compared to persistent carriers. The factors that prevent colonisation in non-carriers are not known but immune responses to
SA are likely to be important. A study of experimental inoculation of healthy volunteers found that nasal infection persisted longer in people who were susceptible to the smallest inoculum. Participants in this study were less susceptible to repeat inoculation with the same strain of SA; yet remained susceptible to infection with other strains.

Serological responses to SA antigens were compared in sera from persistent carriers and non-carriers of SA. Anti-TSST-1 antibody did not protect against SA carriage; IgG antibodies to TSST-1 were higher in persistent carriers and were able to neutralise the function of TSST-1. It was postulated that anti-TSST-1 antibodies might be protective against mortality of SA bacteraemia.

6.1.7 Prior colonisation increases the risk of SA disease, but protects against death from SAB

While it is known that some disease caused by SA, such as skin boils, can follow a chronic course characterised by remissions and relapses; there is evidence that prior exposure to SA does reduce disease severity. It is well known that nasal carriage of SA predisposes an individual to peri-operative infection and that there is a limited role for decolonisation to reduce infections. Wertheim et al screened 14,000 non-surgical patients admitted to hospital to compare the rates of hospital-acquired SA bacteraemia in carriers and non-carriers. The relative risk of bacteraemia was three times greater in people identified as carriers at the time of admission. Pulsed-field gel electrophoresis was performed on all of the SA isolates; in 80% of cases of bacteraemia in carriers of SA, the isolate that caused bacteraemia was identical to the isolate that caused nasal colonisation. Nasal carriers of SA had some protection against mortality when they developed bacteraemia caused by their colonising strain of SA; their mortality rate was significantly less than that of non-carriers who developed SA bacteraemia [SAB].

6.1.8 Antibodies against SA antigens are ubiquitous in adults and develop in infancy

Antibodies have been demonstrated in human sera against almost all known SA virulence factors. Several studies have been performed that measured serological responses to a number of SA antigens in people with infection caused by SA and in healthy non-infected people. These studies all share the findings that antibodies to SA antigens are present in all adults, often in high titre; and that individuals have variation in the antibody titres to different antigens. However, there were no consistent differences identified in the antibody responses between infected and healthy individuals.

For example, Dryla et al found that titres of IgG and IgA antibody against clumping factor B [ClfB] were lower in people with disease caused by SA than in carriers; and higher titres appeared to protect against SA colonisation. In contrast, Verkaik et al found that titres of IgA antibody against ClfB were higher in SA carriers than non-carriers.
Antibodies against SA have also been measured in children. Antibodies to total bacterial lysate can be detected in infants and titres rise with increasing age. This finding is consistent with high rates of exposure to SA in early childhood.

The immune response to SA infection is complex and protective immunity is unlikely to be mediated by antibodies to any single virulence factor. However, it does remain likely that antibodies to secreted toxins, such as TSST-1, are important in reducing the morbidity and mortality of invasive disease caused by SA.

6.1.9 Does immunity to TSST-1 prevent disease?

Immune responses to TSST-1 have been frequently studied over the past three decades.

6.1.10 TSST-1 is a potent superantigen

TSST-1 is classically associated with staphylococcal toxic shock syndrome [TSS], first described in children in 1978. This first case report described the classical features of high fever, profound hypotension, multi-organ failure, diffuse rash and late desquamation in a 15 year old girl with throat and vaginal cultures positive for SA. TSST-1 is a T cell Receptor Vβ2 specific superantigen. Very small quantities can stimulate up to 20% of circulating T cells.

Superantigens bind to major histocompatibility complex class II [MHCII] molecules on antigen presenting cells outside of the antigen-binding groove. They bind simultaneously with the variable region of the T cell receptor, leading to T cell stimulation independent of classical antigen presentation. This leads to the production and release of large amounts of cytokines, such as IL-2, INF-γ and TNFα, which are likely to contribute to the clinical manifestations of TSS. TSST-1 is a member of a larger family of staphylococcal and streptococcal superantigens, some of which are substantially more potent as T cell mitogens in vitro, yet surprisingly staphylococcal TSS only occurs in people infected with tsst+ staphylococcal strains. This suggests that TSS is a result of the amount of TSST-1 that is elaborated in the absence of neutralising anti-TSST-1 antibodies, rather than the specific activity of TSST-1 itself.

6.1.11 Low anti-TSST-1 antibody levels were more common in people who developed TSS

There is a long history of research regarding immune responses to TSST-1, since the increased recognition of cases of TSS related to menses in the 1980s. In 1981, Bergdoll et al reported that high anti-TSST-1 antibody titres were more likely to be detected in the sera of healthy volunteers, than in the acute sera of women who developed TSS. The authors also found that the antibody titre in most cases of TSS did not increase during convalescence; there were a number of women with TSS who did not develop antibodies to TSST-1 at all.
Shortly after the discovery of TSST-1, it was recognised that occasional women were prone to recurrent episodes of TSS with subsequent menses. When acute sera from such women were analysed, nine out of ten had levels of anti-TSST-1 antibodies that were low or not detectable. Within 20 days of the TSS episode, seroconversion was present in approximately half; one-third of women had not seroconverted after one year. A lack of anti-TSST-1 antibodies in women has been associated with younger age and, in North America, with African-American ethnicity.

A murine model also provides evidence that anti-TSST-1 antibodies prevent disease. Mice that had received intraperitoneal vaccination with TSST-1 were more resistant to death from subsequent systemic TSST-1 challenge, than mice that were not vaccinated.

6.1.12 The role of TSST-1 in SAB

The role of TSST-1 in SAB was investigated in two studies performed on the same patient population. The investigators first measured the serum titres of anti-TSST-1 antibody in 103 healthy control subjects. They found that approximately 70% of the control subjects had high titres (greater than 1:100) of anti-TSST-1 antibody. Using a stringent definition of a positive anti-TSST-1 response either as a titre greater than the 95th percentile of the control subjects [1:1300] or as a greater than two-fold rise in titre in convalescence, they then investigated the anti-TSST-1 responses in 63 patients with SAB. 11/12 [92%] who had bacteraemia caused by a TSST-1 producing strain had a positive anti-TSST-1 response. 8/51 [16%] of patients who had SAB caused by a strain that did not produce TSST-1 had a positive antibody response.

In the next study of the same patients with SAB, the authors showed that levels of TNF\(_{\alpha}\) were higher in people with SAB caused by strains that produced TSST-1 and/or other toxins [staphylococcal enterotoxins A, B, C, or D]. Furthermore, patients with high levels of TNF\(_{\alpha}\) were more likely to develop high titres of antibodies against TSST-1 or other toxins in convalescence. Patients, who had high titres of antibodies in acute sera, developed lower levels of TNF\(_{\alpha}\) in response to bacteraemia. These findings provide indirect evidence that high anti-TSST-1 antibody titres might inhibit the detrimental effects of TSST-1 in SAB.

6.1.13 Staphylococcal superantigen-like proteins

The first staphylococcal superantigen-like protein [SSL] to be discovered was SSL-7 [formerly staphylococcal enterotoxin like protein -1]. Other SSLs were discovered by a search of published SA genomes for nucleotide sequences similar to conserved regions of known staphylococcal and streptococcal superantigens. There are 14 known SSLs, 11 of these are found on a pathogenicity island and are numbered according to the order that the genes were present in the nucleotide alignment of published SA genomes. The SSL proteins share structural homology with TSST-1, but do not display superantigen activity.
SSL7 was identified in 100% of SA strains obtained from asymptomatic carriers and patients with bacteraemia\(^{222}\). SSL7 is able to bind both IgA and complement factor 5 [C5] at nanomolar concentrations and prevent monomeric serum IgA binding to the IgA receptor [Fc\(\alpha\)R1] on myeloid cells, such as neutrophils\(^{175}\). SSL7 also prevented complement mediated haemolysis and was shown to significantly improve the survival of *E. coli* cells incubated in human serum.

Resolution of the crystal structure of SSL7 in complex with the heavy chain of IgA has identified the important binding sites, which block the binding of IgA to the Fc\(\alpha\)R1 receptor\(^{223}\). A number of C5 binding sites have also been identified, which allow for a comparison of functional sites in alleles of wild type SA\(^{224}\). A crystal structure of SSL7 bound to complement C5 reveals that SSL7 binds to the MG2 and MG5 domains well away from the C5a domain.

The function of SSL7 is shown in Figure 6.1. Laursen et al proposed that SSL7 forms a pentamer of IgA Fc-SSL7\(_2\)-C5\(_2\), which prevents binding of the C5 convertase and the formation of C5a and C5b\(^{174}\). SSL7 also binds C5b, independently of IgA Fc, and thus, prevents the formation of the membrane attack complex.

---

Figure 6.1 The function of SSL7; a pentamer of IgA Fc-SSL7\(_2\)-C5\(_2\) prevents binding of C5 to C5 convertase; SSL7 also binds C5b to prevent formation of the membrane attack complex.
6.1.15 Immune responses to SSL7

There is only one report of anti-SSL7 antibodies in human sera. This study used ELISA to demonstrate anti-SSL7 antibodies in the sera of 9/10 healthy volunteers at a titre of 1:2000. No studies have measured titres of antibodies to SSL7 in SAB.

6.1.16 The function of SSL9

SSL9 also inhibits complement activity, but does so at a different site to SSL7. SSL9 inhibits the formation of the C3 convertase, critical in the classical and mannose-binding lectin activation pathways. In vitro, SSL9 was shown to inhibit complement mediated haemolysis. It also binds fibronectin and fibrinogen. Binding to an extracellular matrix of fibronectin and fibrinogen might serve to concentrate SSL9 at the site of infection. Anti-SSL9 antibodies were detected in high titre in serum from four healthy volunteers and also from four people with staphylococcal disease.

6.1.17 The function of SSL3 and SSL11

The functions of SSL3 and SSL11 are under investigation. Both belong to a family of SSLs that bind to the trisaccharide sialyllactosamine. SSL3 and SSL11 bind to neutrophils and P-selectin resulting in inhibition of neutrophil rolling along a surface coated with P-selectin (Jackson, 2008; Chung, 2007). SSL3 also binds lactoferrin, which is found in mucosal secretions and has bactericidal activity.

High titres of anti-SSL3 and anti-SSL11 antibody were found in serum from four people with SA infection and also in serum from healthy volunteers. The healthy volunteers tested for antibody to SSL3, SSL9 and SSL11 were all researchers in a laboratory that investigated SA; their high titres of antibodies might have been a consequence of workplace exposure.
6.2 Aims

To measure the concentrations of anti-TSST-1 and anti-SSL antibody levels in acute and convalescent sera from 148 cases of SAB, and in sera from 21 healthy control participants who were not healthcare workers, did not have a history of SA infection and who did not work in a medical research laboratory.

To test these sera for the presence of neutralising antibodies to TSST-1 and SSL7 function.

To test for correlations between the concentration of antibody to SSL3, SSL7, SSL9, SSL11 and TSST-1. This comparison will determine whether antibodies are likely to be directed at epitopes that are conserved among SSL proteins, and whether some individuals lack the ability to mount affective immune responses to all, or most, of these antigens.

Finally, the demographic and clinical features of cases with SAB will be examined in relation to their antibody concentrations, to determine if susceptibility to SA disease is correlated with variation in the immune response to TSST-1 or SSL proteins.
6.3 Materials and Methods

6.3.1 Serum samples

Serological responses to SA antigens were tested using acute and convalescent samples. Acute samples were taken at the time of the first positive blood culture. Most cases of SAB had serum collected at the time of blood culture and the testing laboratory routinely stored any residual sample. This sample was obtained from the laboratory after recruitment to the study. In instances when serum had not been collected at the time of bacteraemia, the sample that was available closest to that time was obtained; or the investigator obtained a blood sample when consent was obtained. All acute serum samples were obtained within 48 hours of collection of the first positive blood culture.

Convalescent samples were collected more than 21 days after the collection of the first positive blood culture. Cases that were discharged from hospital within that time period had samples collected at the community-testing laboratory or in their own home by the principal investigator.

6.3.2 Storage of serum

Each sample was centrifuged at 3500 rpm for 5 minutes at 4°C, and 500uL aliquots of serum/plasma were removed by pipette and stored at -20°C. For each assay, an aliquot was thawed on ice and cryoprecipitate was removed by centrifugation at 12000 rpm and 4°C for 5 minutes. Antibody assays for each antigen were performed in duplicate or triplicate using the same aliquot of serum.

6.3.3 Recombinant SSL proteins

Recombinant SSL9 and SSL3 were provided by Dr Nicola Jackson [University of Auckland]; SSL11 by Dr Matthew Chung [University of Auckland]; and SSL5 by Dr Ries Langley [University of Auckland] \(^ {226-228} \). The NH2-terminal domain of SSL3 is subject to degradation, and a truncated SSL3 protein, which was more stable in solution, was used in the antibody assays. The truncated protein had greater activity in a number of assays when compared to the full-length protein \(^ {226} \).

Antibody responses against three alleles of SSL7 and a single allele of TSST-1 were also determined. The GL1 and GL10 SSL7 alleles [SSL7\(_{GL1}\) and SSL7\(_{GL10}\)] used were derived from genetically diverse strains of SA isolated at Greenlane Hospital in Auckland, and had been cloned as a fusion protein with thioredoxin into the lac operon of \textit{E. coli} AD494 by Dr Ries Langley \(^ {175} \). Two further alleles were expressed and purified: SSL7\(_{4427}\) and SSL7\(_{6301}\); both were cloned by Dr Natasha Gordon [University of Auckland] \(^ {224} \). The genetic diversity of the ssl7 gene is shown in Figure 6.2. Recombinant SSL7\(_{NU4089}\) had also been cloned but was not tested in this study, because of its similarity to SSL7\(_{GL1}\).
Dr Roger Tiedemann [University of Auckland] cloned the TSST-1 allele into *E. coli* as a fusion protein with glutathione-S-transferase [GST] \(^ {229} \). These proteins were expressed and purified as follows.

### 6.3.4 Expression of recombinant protein

*E. coli* cells stored in -80°C were grown overnight in 100ml of shaken LB broth containing ampicillin 50ug/ml\(^ {1} \), chloramphenicol 15ug/ml\(^ {1} \), and kanamycin 35ug/ml\(^ {1} \). This culture was added to 900ml of LB containing the same concentration of antibiotics. Expression of the recombinant protein was induced by the addition of 0.1mM Isopropyl \(\beta\)-D-1-thiogalactopyranoside [IPTG], shaken at 30°C for 4 hours. The cells were pelleted by centrifugation at 5,000 rpm for 20 minutes.

*E. coli* expressing SSL7 alleles were resuspended in 50ml of NTA-I buffer [300mM NaCl, 50mM NaHPO\(_4\), 10mM imidazole, 10% v/v glycerol, pH 8.0] containing 1% Triton X and 100uL 0.1mM phenylmethanesulphonyl fluoride [PMSF].

*E. coli* expressing TSST-1 were resuspended in 50ml GSH buffer [25mM Tris 7.4, 50mM NaCl, 1mM ethylenediaminetetra-acetic acid [EDTA]] containing 1% Triton X and 100uL 0.1mM PMSF.

The resuspended cells were then lysed by sonication [Misonix XL2015]; and the lysate was collected after centrifugation at 10000 rpm for 10 minutes.
6.3.5 Purification of recombinant SSL7 alleles

Each lysate contained soluble SSL7 fused to thioredoxin amongst other cellular debris. Purification was performed by passage over a Nickel sulphate iminodiacetic acid (IDA) sepharose column, which bound the thioredoxin: SSL7 fusion protein. The column contained 3ml of IDA sepharose and cellular debris were washed away with 20 column volumes [60ml] of NTAI buffer. The fusion protein was eluted from the column using NTA-II buffer [300mM NaCl, 100mM imidazole, 50mM NaHPO₄, pH 8.0] and the fusion proteins were cleaved using 2mM dithiothritol (DTT) and 1% [20ug] 3C protease overnight at 4°C.

Imidazole was removed by dialysis into NTA-III buffer [300mM NaCl, 50mM NaHPO₄, pH 8.0], to allow for separation of SSL7 from thioredoxin as they were passed through the IDA Sepharose column. SSL7 passed through the column while thioredoxin [with a polyhistidine tag] remained bound. A sample from each step of the purification procedure was checked by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The final purified protein was dialysed into phosphate buffered saline (PBS) 7.4 [8gL⁻¹ NaCl, 200mgL⁻¹ KH₂PO₄, 1.16gL⁻¹ Na₂HPO₄, 200mgL⁻¹ KCl, pH 7.4], filter sterilised and stored at a final concentration of 1mgml⁻¹.

6.3.6 Purification of recombinant TSST-1

The purification of TSST-1 followed a similar process to SSL7. Glutathione-SH agarose was used to bind GST. After cellular debris was removed the fusion protein was cleaved using 0.2ugml⁻¹ L-1-tosylamido-2-phenylethyl chloromethyl ketone treated trypsin at 37°C for 2 hours. 0.1mM PMSF was added and dialysed into CM I buffer [10mM NaHPO₄, pH6.0] for cation-exchange chromatography to remove the GST.

The cleaved proteins were passed over a column containing carboxymethyl [CM] sepharose [GE Healthcare, Auckland], which bound TSST-1. GST was removed by washing with 20 column volumes [60ml] of CM I buffer. TSST-1 was eluted from the column using 10 column volumes of CM II buffer [50mM NaHPO₄, pH 6.8]. The final protein was dialysed into PBS 7.4 and stored after filter sterilisation at a concentration of 1mgml⁻¹.

6.3.7 Radio-labelling of recombinant proteins

Antibody assays were performed by radio-immunoassay. Each recombinant protein was labelled with radioactive iodine [I¹²⁵] in a fumehood within a radiation suite. 0.5mCi of I¹²⁵ was added to 10uL of 1mgml⁻¹ protein in 0.5M NaPO₄ buffer [13.12gL⁻¹ NaH₂PO₄, 108.6gL⁻¹ Na₂HPO₄]. Coupling occurred during oxidisation of the protein by the addition of 1mgml⁻¹ chloramine T for 60 seconds. The reaction was stopped by the addition of 5mgml⁻¹ of sodium metabisulphite. The radio-labelled protein was separated from free I¹²⁵ by passage over a Sapheaded G-25 [GE Healthcare, Auckland, NZ] size exclusion column. The column was pre-saturated by 10% bovine
haemoglobin in 0.5M NaPO₄ buffer with 1% Triton X to ensure that labelled protein was not lost due to non-specific binding to the column. Each 50uL aliquot that passed through the column was collected and the radioactivity estimated using a Geiger counter. The first radioactive peak to pass through the column represented the labelled protein and the second peak represented the free I¹²⁵. The labelled protein was stored in a lead container at 4°C for use within 5 days.

6.3.8 Preparation of Cowan I S. aureus cells as a protein A source

Due to the large number of samples a radioimmunoassay with a high throughput was required. Cowan I cells were used as a source of protein A to bind heavy chain of IgG. Cowan I cells stored at -80°C were grown overnight in 1L of Luria-Bertani [LB] broth [1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl]. The culture was centrifuged at 5000 rpm for 20 minutes and the supernatant discarded. The cells were resuspended in 0.02% NaN₃ in PBS 7.4 and then centrifuged again. The pellet was resuspended as a 10% solution in 1.5% Formaldehyde in PBS 7.4, and gently mixed for 90 minutes. The cells were washed again in PBS 7.4 and then heat killed at 80°C for 5 minutes. After a final wash in PBS 7.4 the cells were stored at 4°C as a 10% solution in 0.05%Tween 20 PBS 7.4. The final solution was adjusted to an optical density of 2.6 at 540nm for each assay. A single preparation provided enough Cowan I cells for radioimmunoassay of the entire test samples for three antigens.

6.3.9 Radioimmunoassay for antibodies to staphylococcal toxins

Quantification of antibodies to staphylococcal toxins was performed using established methods [Figure 6.3]. 50uL of 10% test serum was added to radio-labelled antigen [to provide 75,000 counts per minute [cpm] per tube] in 100uL of binding buffer [0.5% bovine serum albumin, 0.05% Tween 20 in PBS 7.4]. After 15 minutes at room temperature [RT], antibodies were bound to protein A using Cowan I S. aureus cells as a protein A source. 100uL of Cowan I cells were added to provide an excess of protein A and incubated at RT for 15 minutes. The Cowan cell: antibody: radio-labelled antigen complexes were pelleted by centrifugation at 2500 rpm for 10 minutes. The supernatant was discarded and the pellet was resuspended in wash buffer [0.5% NP40 and 0.05% Tween 20 in PBS 7.4]. The cells were pelleted again and resuspended in wash buffer three times to remove non-specific antibody binding. After the final centrifugation the tube was blotted onto paper to remove any residual unbound labelled protein. The counts per minute [cpm] of each tube provided quantification of the labelled antigen: antibody complexes. Each test sample was measured in three replicate assays, performed on the same day. Each assay included negative controls of 50uL 10% fetal calf serum [FCS] and a positive control of 50uL 1:1000 polyclonal purified rabbit IgG antibody specific for each antigen. A negative result was one that fell within 3 standard deviations of the mean results of replicate assays performed using 10% FCS negative controls.
Figure 6.3 Schematic of quantification of antibody by radioimmunoassay using Cowan I cells as a source of protein A.

(A) specific (red) and non-specific (blue) antibodies bind to radio-labelled antigen (yellow star)

(B) protein A on the surface of Cowan I S.aureus cells binds antibody

C) non-specific antibody is removed after three washes, and the total gamma emission (counts per minute) of antigen bound to specific antibody is measured
6.3.10 Preparation of PBMCs for a T cell proliferation assay

The presence of functional anti-TSST-1 antibodies was determined by inhibition of TSST-1 mediated T cell proliferation. T cells were prepared from whole blood obtained from a healthy volunteer. The blood was diluted 1:1 in sterile PBS 7.4 and passed gently over a Ficoll-paque gradient and centrifuged at 750g for 15 minutes. Peripheral blood mononuclear cells [PBMC] were removed by pipette and were washed twice in sterile PBS 7.4 followed by centrifugation at 1200 rpm for 7 minutes. The cells were resuspended at 37°C in cRPMI [Roswell Park Memorial Institute 1640 media [RPMI] containing 10% fetal calf serum, 100 U mL^{-1} penicillin, 100 µg mL^{-1} streptomycin, 2 mM L-glutamine, 1% pyruvate].

6.3.11 T cell proliferation assay

100uL of PBMCs adjusted to a concentration of 10^6 cells ml^{-1} was added to a 96 well round bottom plate with serial dilutions of TSST-1 starting at 200ng ml^{-1} in 100uL cRPMI. 50ng ml^{-1} of streptococcal superantigen SMEZ-2 was used as a positive control. The PBMCs were incubated for 72 hours in 5%CO_2 for 72 hours. Tritiated thymidine [^{3}H-TdR] was incorporated into proliferating T cells after 0.25uCi was added to each well and incubated at 37°C for a further 16 hours.

T cells were harvested onto a filter paper mat using a Tomtec cell harvester [Wallac OY, Turku, Finland]. The mat was then sealed in a plastic sleeve containing enough scintillant to cover the mat evenly. β emissions from ^{3}H-TdR incorporated into proliferating T cells were then counted using a β-counter [Wallac OY, Turku, Finland]. The lowest concentration of TSST-1 that caused maximal T cell proliferation [15ng ml^{-1}] was used in the functional anti-TSST-1 antibody assay [Figure 6.4].

![Figure 6.4 T cell proliferation [counts per minute, cpm] mediated by TSST-1.](image)
6.3.12 Functional anti-TSST-1 antibody assay

The assay to detect functional antibody against TSST-1 was performed in the same manner as the T cell proliferation assay. 20µL of neat human serum was added to each well, to give a final concentration of approximately 10%. Each well contained 100µL of PBMCs suspended 10^6 ml^-1 in cRPMI and 100µl of TSST-1 [final concentration of 15ng ml^-1] in cRPMI.

Each assay was performed in triplicate and the results were expressed as a percentage of inhibition by dividing the degree of T cell proliferation [β-count] of each well by the T cell proliferation in control wells that did not contain human serum. 10% fetal calf serum was used as a negative control.

6.3.13 ELISA of SSL7 binding to complement factor C5 in human serum

A sandwich enzyme-linked immunosorbent assay [ELISA] was performed to test for SSL7 binding to C5. A 96 well ELISA plate was coated with 50µL per well of 1mgml^-1 SSL7_GL10, and then washed with PBS-T [PBS 7.4 with 0.05% Tween 20]. The SSL7_GL10 allele was used, because it was the most potent inhibitor of alternative complement activation pathway mediated haemolysis of sheep red blood cells [see 9.3.15] 224.

50µL of 10% human serum was added and incubated at ambient temperature for 30 minutes. After washing in PBS-T, 1:2000 rabbit anti-human C5 antibody [>>>] was added to detect any C5 in the sera that had bound to SSL7_GL9. Another wash was performed and C5 bound to SSL7_GL10 was detected by the addition 1:2000 anti-rabbit IgG1 antibody conjugated to horseradish peroxidase [HRP]. 50µL of developing solution was added to each well [50 mM citric acid, 100 mM Na2HPO4, 0.012% H2O2, and 0.5 mg ml^-1 o-phenylenediamine dihydrochloride] in the dark for 5 minutes. The developing reaction was stopped by the addition of 50µL 10% HCl. The optical density at 490nm [OD_490] was recorded using a µQuant spectrometer [BioTek Instruments Inc, Vermont, USA].

A number of negative controls were tested. C5 was not detected after serum was added to wells coated with TSST-1, SSL3 and SSL11. Other complement factors could not be detected by anti-C1q, anti-C3, and anti-C4 antibodies. A monoclonal antibody against the formed membrane attack complex [C5-C9] was weakly positive. The addition of supplemental C5 or IgA did not add to the sensitivity of the assay.

An ELISA was performed to determine if complement factor C5 in stored serum samples bound to SSL7_GL9. Serum from two healthy donors that had been stored at -20°C for more than one month all contained enough C5 to be detected at low dilutions; no C5 was detected in fetal calf serum [Figure 6.5].
6.3.14 Detecting anti-SSL7 antibodies that reduce binding of SSL7 to complement factor C5

ELISA was used to test sera from cases with SAB and from healthy controls for binding of C5 to SSL7\textsubscript{GL9}. Each serum sample was tested in triplicate; and negative sera gave mean OD\textsubscript{490} values that were not more than 3 standard deviations greater than the mean results from fetal calf serum negative controls. Negative samples either had C5 concentrations below the sensitivity of the assay, or had anti-SSL7 antibodies that prevented the binding of C5 to SSL7. Thus, an ELISA was used to detect C5 in the samples that tested negative for C5 binding to SSL7.

Serial dilutions of serum were added to an ELISA plate and incubated at room temperature for 1 hour. Rabbit anti-C5 IgG [1\text{mg ml}^{-1}, 1:4000] was used to bind to captured C5 and the plate was developed and read after the addition of mouse anti-rabbit IgG antibody conjugated to HRP. Fetal calf serum was used as a negative control and sera from two healthy volunteers were used as positive controls. A positive result was indicated by the presence of C5 in a dilution of serum that was within two dilutions of the negative values for the positive controls. All of the sera detected C5 at dilutions of 1:320 and 1:640, which were the dilutions before the binding of C5 began to fall in the healthy control samples [Figure 6.5]. Thus, a negative result indicated the presence of antibody that interfered with C5 binding [Figure 6.6].
Figure 6.6 Schematic of ELISA to detect antibody inhibition of C5 binding to SSL7.
6.3.15 SSL7 alternative pathway complement inhibition assay

SSL7 binds to C5 and inhibits the formation of the membrane attack complex; SSL7 has been shown to inhibit complement mediated haemolysis. An assay that measured the inhibition of alternative pathway complement mediated haemolysis of sheep red blood cells by SSL7 has been developed.

Sheep red blood cells were washed several times in GHE buffer (GH buffer: 75%v/v 80mM HEPES, 280mM NaCl, 0.15% porcine gelatin and 1mM Ethylene glycol tetra-acetic acid [EGTA] and 10mM MgCl₂) and centrifuged at 1250 x g for 5 minutes until the supernatant was clear. EGTA was used to chelate the calcium ions required for the classical complement activation pathway. Thus, haemolysis was not antibody dependant and the RBC did not need to be sensitised. 50uL of sheep RBC resuspended in GHE buffer at 2% were added to 50uL of fresh human serum and 5uL of GHE buffer in a 96 well plate.

Haemolysis was inhibited by serial dilutions of 50uL SSL7GL10, starting at 16uM per well, to determine the appropriate concentration of SSL7 to use in a functional antibody assay [Figure 6.7]. The plate was incubated on a shaker for 30 minutes at 37°C and was then centrifuged at 4500 rpm for 5 minutes to separate intact sheep red blood cells. 50uL of supernatant was removed from each well and placed onto a new 96 well plate to record the amount of haemoglobin pigment in each well. The optical density at 412nm [OD₄₁₂] for each well was recorded using a uQuant spectrometer. The percentage of total haemolysis was determined by dividing the degree of haemolysis [OD₄₁₂] for each well by the degree of haemolysis of the control wells that did not contain SSL7.

![Figure 6.7 SSL7 inhibition of haemolysis of sheep RBC by human serum](image)
6.3.16 Assay for anti-SSL7 antibodies that neutralise SSL7 inhibition of complement mediated haemolysis

The SSL7 alternative pathway complement inhibition assay was modified to detect the presence of anti-SSL7 antibodies that neutralised SSL7 inhibition of haemolysis. 5uL of test serum was added to each well in place of the 5uL aliquot of GHE buffer, to give a final concentration of approximately 10% test serum. Each assay was performed in three replicates. Any residual complement activity of the test samples was removed by prior incubation at 56°C for 15 minutes. 10% fetal calf serum was used as a negative control in each replicate of the assay [Figure 6.8]. The background haemolysis [OD$_{412}$] was measured in 50uL of supernatant from wells that contained only GHE buffer and sheep red blood cells, and was subtracted from each result. OD readings were obtained prior to the addition of sheep RBC and were subtracted from the readings obtained from the assay, to account for variance in pigmentation between the serum samples.

Figure 6.8 Fetal calf serum [FCS] does not restore haemolysis of sheep RBC inhibited by SSL7
The final result was expressed as the haemolysis ratio, which was calculated by dividing the mean haemolysis of each test sample [minus 2 standard deviations] by the mean haemolysis of the fetal calf serum controls [plus 2 standard deviations]. A result of 1 or more indicated that some haemolysis had been restored by antibody binding to SSL7.

The error of the assay was dependant on the concentration of SSL7GL10 used. An initial assay of the test sera was performed using a concentration of 0.4uM, which was the lowest concentration that completely inhibited haemolysis [Figure 6.7]. The results that were obtained contained a large amount of error caused by low background levels of haemolysis and so the concentration of SSL7GL10 used in the final assay was increased to 0.8uM. This resulted in a marked reduction in the sensitivity of the assay.
6.4 Results

Serum samples were obtained from 148 people with bacteraemia caused by SA. 150 cases of SAB were admitted to Auckland City Hospital or Middlemore Hospital and were recruited to the study; however, serum was not obtained from two participants within 48 hours of their SAB. Serum was also collected from 21 healthy household contacts of these cases. These contacts were not healthcare workers and none had been admitted to hospital in the preceding year. The demographic features of the cases of SAB were described in detail in Chapter 1. The demographic features of the participants whose immune responses to SA were tested are shown in Table 6.1.

<table>
<thead>
<tr>
<th>sample [n]</th>
<th>mean age, s.d.</th>
<th>female</th>
<th>Māori</th>
<th>Pacific people</th>
<th>Other ethnicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteraemia - Auckland City Hospital [94]</td>
<td>57, 20 years</td>
<td>31 [32%]</td>
<td>16 [17%]</td>
<td>22 [23%]</td>
<td>56 [60%]</td>
</tr>
<tr>
<td>Bacteraemia - Middlemore Hospital [54]</td>
<td>53, 20 years</td>
<td>17 [31%]</td>
<td>16 [30%]</td>
<td>22 [40%]</td>
<td>16 [30%]</td>
</tr>
<tr>
<td>Healthy control [21]</td>
<td>49, 20 years</td>
<td>11 [52%]</td>
<td>2 [10%]</td>
<td>7 [33%]</td>
<td>12 [57%]</td>
</tr>
</tbody>
</table>

Table 6.1 Demographic features of the participants whose immune responses to SA were tested
6.4.1 Convalescent serum

Acute serum samples were obtained from 148 participants. Convalescent samples were those provided more than 21 days after the date of collection of the first positive blood culture. 16 participants died within this time period. 88/132 [67%] of the remaining study participants provided a convalescent sample. The convalescent samples were collected a mean of 30 days, s.d. 13 days, after the date of collection of the first positive blood culture [Figure 6.9].

![Figure 6.9 The time [days] of collection of 88 convalescent serum samples.](image-url)
6.4.2 Antibodies against TSST-1

The mean anti-TSST-1 antibody concentration, expressed as counts per minute of radio-labelled recombinant TSST-1 bound to antibody is shown in Figure 6.10. The mean antibody concentrations were compared between acute, convalescent and healthy sera. Levene’s test of homogeneity of variance indicated that group variances were not equivalent. However, the group variances were equivalent after the data had been transformed logarithmically [Levene’s test $p=0.3$]. Thus ANOVA was performed on the transformed data, and did not find a significant difference between acute and convalescent sera. Likewise there was no difference between the mean anti-TSST-1 antibody concentration in healthy people and cases of SAB.

![Box plot showing mean anti-TSST-1 antibody concentration](image)

Figure 6.10 The mean anti-TSST-1 antibody concentration [counts per minute, cpm] of acute [$n=148$] and convalescent [$n=88$] sera from cases of SAB and from healthy controls [$n=21$]. Also shown are the mean positive and negative control samples. Each box shows the range [the tips of the vertical lines from each box], the median [central horizontal line] and the 25th and 75th percentiles [interquartile range, grey box] for each sample type. The circles indicate outlier results that are more than 1.5 times greater or less than the interquartile range.

6.4.3 No increase in anti-TSST-1 antibody level after bacteraemia

The paired acute and convalescent anti-TSST-1 antibody concentrations in 88 cases of bacteraemia are shown in Figure 6.11. A negative value was one that fell within three standard
deviations of the mean value of the negative controls. Two patients seroconverted from a negative to a positive value, although in only one case was there a dramatic increase in antibody level. Antibody concentrations increased in a small number of patients during convalescence; however, sera from the majority of patients demonstrated no overt change in level between acute and convalescent sera. In several cases there was a large drop in antibody level between acute and convalescent samples.

![Graph showing antibody concentrations in paired acute and convalescent sera.](image)

**Figure 6.11** Anti-TSST-1 antibody concentrations in paired acute and convalescent sera. The dashed line indicates the negative cut-off.

### 6.4.4 A high proportion of people with bacteraemia have antibodies that inhibit the function of TSST-1

The mean percentage of inhibition of TSST-1 mediated T cell proliferation by sera from cases of SAB and from healthy controls is shown in Figure 6.12. Sera from 134/148 [86%] cases with bacteraemia contained sufficient anti-TSST-1 antibody in their acute serum sample to inhibit more than 25% of the total T cell proliferation. Sera from 17/21 [81%] of healthy controls contained sufficient anti-TSST-1 antibody in their acute serum sample to inhibit more than 25% of the total T cell proliferation. There was no significant difference between the two groups. The proportion of the cases of SAB whose acute serum inhibited the function of TSST-1 almost completely [>90% inhibition of T cell proliferation] was not different to the proportion of healthy controls whose serum almost completely inhibited the function of TSST-1 [cases of SAB; 86/148, 58%; healthy controls 14/21, 67%].

---

181
Figure 6.12 The mean percentage of inhibition of TSST-1 mediated T cell proliferation by sera from cases of SAB and sera from healthy controls, and negative controls [50% fetal calf serum, FCS]
6.4.5 The concentration of anti-TSST-1 antibody correlates with neutralisation of TSST-1 function

There was a significant positive correlation between the anti-TSST-1 antibody concentration and the percentage of inhibition of T cell proliferation for the 148 acute sera from cases with SAB and 21 healthy control samples [Pearson’s coefficient 0.517, p<0.01] [Figure 6.13].

Figure 6.13 The correlation between anti-TSST-1 antibody concentration and the percentage of inhibition of T cell proliferation in acute sera from 148 cases of SAB. The rectangle indicates individuals whose sera efficiently neutralised the function of TSST-1. The linear line of best fit and 95% confidence intervals [dashed lines] are shown.
6.4.6 Anti-SSL7 antibodies are ubiquitous in adults

Anti-SSL7 antibodies were detected in all of the sera tested. When antibodies were tested against the SSL7_{4427} allele, all serum samples had anti-SSL7 antibody concentrations that were more than ten times the mean value obtained from the negative controls [10% fetal calf serum]. There was no significant difference in the mean serum antibody values between healthy control participants and cases with SAB [Figure 6.14].

![Box plot showing mean anti-SSL7_{4427} antibody concentrations](image)

Figure 6.14 The mean anti-SSL7_{4427} antibody concentrations in acute [n=148] and convalescent [n=88] sera from cases of SAB and from healthy controls [n=21].
6.4.7 No important difference between anti-SSL7 antibody concentrations in acute and convalescent sera

A statistically significant difference was detected between the mean values of acute and convalescent sera [ANOVA of log [mean cpm], F=17.1, 1 d.f., p<0.001]. This difference represented a decrease in mean serum antibody values during convalescence [Figure 6.15].

Figure 6.15 Anti-SSL7 antibody concentrations in paired acute and convalescent sera. The dashed line indicates the negative cut-off.
6.4.8 No difference in anti-SSL7 antibody values between different alleles of SSL7

Antibody concentrations against four different alleles of SSL7 were measured. These alleles were chosen because they were representative of the diversity in the SSL7 amino acid sequence [see Figure 6.2]. There was a particularly strong correlation between antibody concentrations against the SSL74427 and SSL7GL1 alleles [Figure 6.16], and there was a strong positive correlation between all four of the alleles [Table 6.2].

Figure 6.16 Correlation between anti-SSL74427 and anti-SSL7GL1 antibody concentrations. The linear line of best fit and 95% confidence intervals [dashed lines] are shown.
<table>
<thead>
<tr>
<th>SSL7&lt;sub&gt;4427&lt;/sub&gt;</th>
<th>SSL7&lt;sub&gt;GL1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>0.54</td>
<td>0.56</td>
</tr>
<tr>
<td>0.61</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>SSL7&lt;sub&gt;GL10&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Table 6.2 Pearson’s coefficient for the correlation of antibody concentrations between different alleles of SSL7. All correlations were statistically significant [p<0.001].

6.4.9 Anti-SSL7 antibodies that block the binding of SSL7 to C5 are uncommon

15/169 [9%] of the acute sera from cases of SAB and from healthy controls had antibodies that blocked the direct binding of SSL7<sub>GL10</sub> to C5; in these samples the mean optical density [OD] of the triplicate assays at 490nm [OD<sub>490</sub>] was less than 0.1 [within 3 standard deviations of the results from the fetal calf serum negative controls]. Anti-SSL7 antibodies that blocked binding to C5 were more prevalent in sera from healthy controls [5/21, 24%] than in sera from cases of SAB [10/148, 7%] [Fisher’s exact test, p=0.02]. There was no significant correlation between blocking of C5 binding and the concentrations of anti-SSL7<sub>GL10</sub> antibodies determined by the radio-immunoassay. That is, high anti-SSL7 titres did not correlate with blocking activity.

When the definition of a positive anti-SSL7 blocking antibody result was extended to include the 25<sup>th</sup> percentile of mean OD<sub>490</sub> values [the lowest 42 results], there was no difference between sera from acute cases of SAB [37/148, 25%] and sera from healthy control subjects [5/21, 24%].
6.4.10 Anti-SSL7 antibodies that neutralise SSL7 mediated inhibition of haemolysis are uncommon

94 acute serum samples from cases of SAB at Auckland City Hospital were tested for anti-SSL7 antibodies that neutralised SSL7<sub>GL10</sub> mediated inhibition of haemolysis. The results were then correlated with the presence of anti-SSL7 antibodies that blocked the binding of SSL7<sub>GL10</sub> to C5.

9/94 [10%] of the sera caused haemolysis to be restored to above the level found with the fetal calf serum negative controls [Figure 6.17]. The level of haemolysis [OD<sub>412</sub>] in these samples was greater than the mean haemolysis of the fetal calf serum negative control; hence the haemolysis ratio was greater than 1. In 5/94 [5%] cases the haemolysis ratio was high; that is, the anti-SSL7 antibody in the sample restored haemolysis to levels more than 2 times the mean value of the negative controls. 4/5 [80%] of these samples contained antibody that blocked binding of immobilized SSL7 to C5; none of the four samples with a haemolysis ratio that was greater than 1 but less than 2 contained anti-SSL7 antibodies that blocked binding of SSL7<sub>GL10</sub> to C5. Two-thirds of the cases [61/94, 65%] had haemolysis ratios less than 0.5.

![Figure 6.17 Neutralisation of haemolysis by functional anti-SSL7 antibodies in acute sera from 94 cases of SAB at Auckland City Hospital. The dashed line indicates the cut-off for positive results.](image)

Figure 6.17 Neutralisation of haemolysis<sup>1</sup> by functional anti-SSL7 antibodies in acute sera from 94 cases of SAB at Auckland City Hospital. The dashed line indicates the cut-off for positive results.

<sup>1</sup>The haemolysis ratio was the ratio of haemolysis in test serum samples [-2 s.d.] to the mean haemolysis in fetal calf serum negative control samples [+2 s.d.].
There was a weak but statistically significant correlation between the haemolysis ratio and the prevention of binding of C5 to immobilized SSL7 [OD$_{412}$] [Pearson’s coefficient = -0.253, p=0.014] [Figure 6.18]. There was no significant correlation between the haemolysis ratio and the level of anti-SSL$^{\text{GL10}}$ antibody measured by radio-immunoassay.

![Graph showing correlation between haemolysis ratio and mean C5 binding to SSL7](image)

**Figure 6.18** The correlation$^1$ between blocking of SSL$^7$ binding to C5 [OD$_{412}$] and neutralisation of SSL$^7$ mediated inhibition of haemolysis [haemolysis ratio] by anti-SSL$^7$ antibodies in sera from 94 cases of SAB.

$^1$Pearson’s coefficient was -0.253 [p=0.014]
6.4.11 Anti-SSL3 antibodies are ubiquitous in adults

Anti-SSL3 antibodies were detected in all of the samples that were tested. There were no differences in the mean antibody concentrations between acute and convalescent sera, however, the mean antibody concentrations were significantly lower in healthy control samples than in acute sera from cases of SAB [ANOVA, F=19.43, 1 d.f., p<0.001] [Figure 6.19].

Figure 6.19 The mean anti-SSL3 antibody concentrations in acute [n=148] and convalescent [n=88] sera from cases of SAB and from healthy controls\(^1\) [n=21].

\(^1\)Healthy vs. acute, ANOVA, F=19.43, 1 d.f., p<0.001
6.4.12 Anti-SSL9 antibodies are ubiquitous in adults

Anti-SSL9 antibodies were also detected in all of the samples tested, but there were no significant differences in the antibody concentrations between acute and convalescent sera; and no differences between sera from cases of SAB and from healthy volunteers [Figure 6.20].

Figure 6.20 The mean anti-SSL9 antibody concentrations in acute [n=148] and convalescent [n=88] sera from cases of SAB and from healthy controls [n=21].
6.4.13 A small proportion of adults lack anti-SSL11 antibodies

12 people did not have anti-SSL11 antibodies in their sera, their radio-immunoassay results were within 3 standard deviations of the results of the fetal calf serum negative controls. There was no difference between the number of cases of SAB, who did not have anti-SSL11 antibodies in acute serum samples [10/148, 7%] and the number of healthy controls [2/21, 10%], who did not have anti-SSL11 antibodies. There was no difference between the mean antibody concentrations of acute and convalescent sera from cases of SAB; and there was no difference between the mean antibody concentrations in acute sera from cases of SAB and healthy controls [Figure 6.21].

Figure 6.21 The mean anti-SSL11 antibody concentrations in acute [n=148] and convalescent [n=88] sera from cases of SAB and from healthy controls [n=21].
6.4.14 Seroconversion to SSL11 does occur during SAB

A convalescent sample was available in seven cases of SAB who did not have anti-SSL11 antibody in their acute samples. 5/7 [71%] of these cases had anti-SSL11 antibody in their convalescent serum samples [Figure 6.22]. In several cases, high concentrations of anti-SSL11 antibody in acute samples fell to low concentrations during convalescence.

![Figure 6.22 Anti-SSL11 antibody concentrations in paired acute and convalescent sera. The dashed line indicates the negative cut-off.](image-url)
6.4.15 Radioimmunoassay for anti-SSL5 antibodies

The fetal calf serum negative control values in the radioimmunoassay for anti-SSL5 antibodies were high and there was only a modest separation between the positive and negative controls. This indicated either the presence of anti-SSL5 antibodies in the negative controls [antibodies against other SSLs were not identified in fetal calf serum] or non-specific binding to immunoglobulin or to the Cowan SA cells themselves. SSL5 is known to bind to sialic acid residues, and the radioimmunoassay was repeated using a T175P SSL5 mutant [SSL5\textsubscript{T175P}] that does not bind to sialic acid \textsuperscript{228}. The radioimmunoassay for anti-SSL5\textsubscript{T175P} antibodies also returned values for the negative controls that were higher than expected. When a small number \([n=20]\) of acute serum samples from cases of SAB were assayed, the results were similar to the negative controls [Figure 6.23]. This assay was not performed on all of the serum samples, because the separation between the values of the positive and negative controls was insufficient to allow informative interpretation and to avoid the unproductive use of the remaining serum samples.

Figure 6.23 Radioimmunoassay for anti-SSL5 antibodies in 20 cases of SAB. The positive control [polyclonal rabbit anti-SSL5 antibody] values and negative control [fetal calf serum] values for the assay of wild type SSL5 and the T175P SSL5 mutant are shown.
6.4.16 Correlation between anti-TSST-1 antibodies and anti-SSL3, SSL7, SSL9 and SSL11 antibodies

The antibody concentrations against staphylococcal toxins [TSST-1, SSL3, SSL7, SSL9 and SSL11] were tested to determine whether individual patient’s antibody concentrations against any of the staphylococcal toxins were correlated with those against other toxins.

There was a significant correlation between concentrations of anti-SSL9 antibody and those of all other antigens, but there were no other significant correlations found [Table 6.3].

<table>
<thead>
<tr>
<th></th>
<th>TSST-1</th>
<th>SSL3</th>
<th>SSL7</th>
<th>SSL9</th>
<th>SSL11</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.079</td>
<td>0.072</td>
<td>0.315†</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.072</td>
<td>0.152</td>
<td>0.594†</td>
<td>0.117</td>
<td>0.138</td>
<td>0.127†</td>
</tr>
</tbody>
</table>

Table 6.3 Pearson’s coefficient for the correlation of antibody concentrations between different SSL proteins and TSST-1.

6.4.17 No correlation between anti-SSL antibodies and the age of cases of SAB

The description of 150 cases of SAB, presented in Chapter 1, indicated that susceptibility to SAB increased with increasing age. The effect of age on the prevalence of SAB was most striking for cases of SAB whose ethnicity was neither Māori nor Pacific. The antibody concentrations against staphylococcal toxins were evaluated to determine whether susceptibility to SAB, with increasing age, was partly caused by a decline in humoral immunity to SA antigens.

There was no relative decline in anti-TSST-1 antibody concentrations with increasing age. In fact, there was a weak positive correlation between anti-TSST-1 antibody values and age [Pearson’s coefficient 0.209, p=0.01]. There was also a weak positive correlation with age and anti-TSST-1 antibody that impaired TSST-1 mediated proliferation of T cells [Pearson’s coefficient 0.167, p=0.042] [Figure 6.24].

There was no significant correlation between the ages of 148 cases of SAB and anti-SSL3, SSL7, SSL9, and SSL11 antibodies.
Figure 6.24 The correlation between the percentage of inhibition of T cell proliferation mediated by TSST-1 and age in 148 cases of SAB. The linear line of best fit is shown, with upper and lower 95% confidence intervals [dashed lines].

Pearson’s coefficient 0.167, p=0.042

6.4.18 Gender has no effect on the development of antibodies to staphylococcal toxins

The mean value of anti-TSST-1 antibodies in 47 female cases of SAB was not different to the mean values in 101 male cases of SAB [10490 compared to 9595, ANOVA F=0.221, p=0.639]. Likewise, there were no significant differences in the mean anti-SSL3, SSL7, SSL9 and SSL11 antibody values between female and male cases of SAB.

6.4.19 Ethnic variation in anti-TSST-1 antibody and anti-SSL9 antibody

The rates of SAB described in Chapter 1 were higher for Māori and Pacific people compared to people of other ethnic groups. Māori and Pacific people with SAB were more deprived than people of other ethnic groups, however, susceptibility to SAB might be multi-factorial and, in part, due to differences in immune response to SA. There were no differences in the anti-SSL3, SSL7 and SSL11 antibody concentrations in acute serum between cases of SAB who were Māori, Pacific people, or people of other ethnic groups.
There was a statistically significant difference between the mean anti-SSL9 antibody concentrations in Māori [49061 cpm] and Pacific cases [46924 cpm] of SAB, compared with cases of other ethnic groups [56869 cpm] [ANOVA, Māori vs. Other F=4.576, p=0.035; Pacific vs. Other =9.386, p=0.003]. There was no difference between the ethnicity of 11 cases of SAB who had anti-SSL9 antibody concentrations that were 1.5 times below the interquartile range [outliers in Figure 6.20]. 3/32 [9%] of these cases were Māori, 3/44 [7%] were Pacific people, and 5/72 [7%] were people of other ethnic groups.

A striking difference was found in anti-TSST-1 antibody concentrations between ethnic groups. The mean anti-TSST-1 antibody concentrations were lower in Māori [7253 cpm] and Pacific people [4992 cpm] with SAB compared with cases of other ethnicities [13926 cpm]. ANOVA, performed on mean cpm values after logarithmic transformation, demonstrated that these results were statistically significant [ANOVA Māori vs. Other, F=14.3, p<0.001; Pacific vs. Other, F=28.7, p<0.001] [Figure 6.25].

---

Figure 6.25 The mean anti-TSST-1 antibody concentration in cases of SAB who were Māori [n=32], Pacific [n=44] and people of other ethnicity [n=72]1.

1 ANOVA Other versus Māori and Other versus Pacific, p<0.001
A high proportion of sera from cases of SAB contained antibody that inhibited T cell proliferation mediated by TSST-1. The mean percentage of inhibition was significantly lower in Māori [70%] and Pacific people [66%] compared to people of other ethnic groups [88%] [ANOVA, Māori vs. Other, F=10.9, p<0.001; Pacific vs. Other F=18.1, p<0.001] [Figure 6.26].

![Figure 6.26 The percentage of inhibition of TSST-1 function in cases of SAB who were Māori [n=32], Pacific people [n=44] and people of other ethnic groups [n=72].](image)

Figure 6.26 The percentage of inhibition of TSST-1 function in cases of SAB who were Māori [n=32], Pacific people [n=44] and people of other ethnic groups [n=72].

1Other versus Māori and Other vs. Pacific, ANOVA p<0.01
6.4.20 No effect of deprivation on antibodies against staphylococcal toxins

When the 148 cases of SAB were grouped according to the deprivation index of their neighbourhood, there were no statistically significant differences in anti-SSL3, SSL7, SSL9, SSL11 and TSST-1 antibody concentrations between cases who lived in the most deprived and least deprived neighbourhoods. There was a trend towards decline in the mean percentage of inhibition of TSST-1 function with increasing deprivation. The mean percentage of inhibition was 85% for cases who lived in the least deprived areas compared with mean values of 81% and 74% for cases from moderately deprived and the most deprived areas respectively [ANOVA, F=1.672, p=0.191] [Figure 6.27].

Figure 6.27 Percentage of inhibition of TSST-1 function by serum from cases of SAB from the least deprived areas [deprivation score category 1], moderately deprived areas [deprivation score category 2] and the most deprived [deprivation score category 3] areas of Auckland.

The differences between groups were not statistically significant [ANOVA, p=0.191].
There were no differences in mean antibody values against staphylococcal toxins or the percentage of inhibition of TSST-1 function for other potential markers of deprivation including annual income under $20,000, the proportion of people receiving a sickness benefit, and cigarette smoking.

The mean inhibition of TSST-1 mediated T cell proliferation was 84% among people who owned their own home compared with 72% for people who did not own their own home [ANOVA, F=5.713, 1 d.f., p=0.018]. This difference was due to the low proportion of Māori and Pacific cases of SAB who owned their own homes [17/76, 22%] compared with the proportion of cases of other ethnicities [54/74, 73%, Fisher’s exact test, p<0.01]. There was no difference in the % of inhibition of TSST-1 function between Māori and Pacific cases who owned their home [63%] or not [68%]; nor was there any difference between cases of other ethnicities who owned their home [89%] or not [82%].

6.4.21 Antibodies against staphylococcal toxins do not prevent death caused by SAB

There were no statistically significant difference in the antibody concentrations or inhibition of TSST-1 function between cases who died from SAB and cases who survived. Likewise, there were no correlations between concentrations of antibodies against the SSL proteins and death caused by SAB. Cases with higher acute physiology and chronic health evaluation [APACHEII] scores were more likely to die from their infection, yet there was no positive correlation between antibody concentrations against staphylococcal toxins and the APACHEII score. A weak inverse correlation between SSL3 antibody level and APACHE II score was detected [Pearson’s coefficient -0.255, p=0.008].

6.4.22 Other demographic factors were not associated with differences in antibodies against staphylococcal toxins

There was no difference in antibody concentrations against staphylococcal toxins or the percentage of inhibition of TSST-1 function between those who acquired their infection in the community or in hospital.

There were no differences in the mean value of anti-TSST-1 antibody level or the percentage of inhibition of TSST-1 function between cases with end stage renal failure and cases who did not have renal failure.

Finally, there were no differences in the antibody concentrations against staphylococcal toxins or the percentage of inhibition of TSST-1 function between cases who had bacterial endocarditis, skin and soft tissue infection, or infection of a peripheral or central venous catheter.
6.5 Discussion

This study demonstrated high concentrations of antibodies against almost all SSL proteins tested and against TSST-1 in healthy adults and adults with SAB. These findings are consistent with previous reports and are indicative of frequent subclinical infection with different SA strains, beginning in the first few months of life.82,101

6.5.1 Immune responses to SSL proteins

The current study was the first to document the serum immune responses to SSL proteins in a large number of cases of SAB and in healthy adults. All of the subjects had high serum concentrations of antibody to SSL3, SSL7 and SSL9. The concentrations of antibodies against SSL11 were more varied, and there were a few cases of SAB and healthy adults who lacked any detectable anti-SSL11 antibody. In a small number of subjects, seroconversion to SSL11 was demonstrated between acute and convalescent serum samples. The SSL proteins share structural homology but there was no evidence of a conserved epitope contributing to the antibody concentrations; a strong correlation was observed between each subject’s antibody concentrations against SSL3 and SSL9, but not between antibody concentrations against other SSL proteins. There was no evidence that some people fail to mount antibody responses against a range of staphylococcal toxins. The findings provide evidence that SSL proteins are commonly expressed during infection and are consistently recognised by the human immune system, supporting their putative roles in host immune evasion and bacterial survival.

6.5.2 Antibodies against TSST-1 correlate with neutralisation of TSST-1 function

It was found that concentrations of anti-TSST-1 antibodies were strongly correlated with inhibition of TSST-1 mediated proliferation of T cells. Highly efficient neutralisation was observed in 60% of the individuals tested. There was no difference in the mean serum anti-TSST-1 antibody concentrations between cases of SAB and healthy volunteers.

A previous study found high titres of anti-TSST-1 antibody in 98% [120/123] of healthy volunteers over the age of 50, but in only 52% [110/210] of people under the age of 15 years.231 In the current study all 21 of the healthy controls had concentrations of anti-TSST-1 antibody that were at least several fold greater than the negative control values. The correlation between each individual’s anti-TSST-1 antibody concentration and age was less marked than that found by Vergeront et al; however, all of the participants in the current study were over the age of 15 years. High titres of anti-TSST-1 antibody do not appear to protect against SAB, but low titres might increase the risk of adverse outcomes following SAB.88 In the current study, a small number of cases with SAB lacked antibody against TSST-1 but there was no increase in mortality or relapse in these cases, and they were not more severely unwell [as determined by their APACHE II score.
at the time of SAB] than cases with higher concentrations of antibody. Likewise, there was no difference in the outcome of SAB between cases that lacked neutralising antibody against TSST-1 [inhibition < 90%] and those who had neutralising antibody against TSST-1.

6.5.3 Antibodies against SSL7 were ubiquitous but antibodies the blocked binding of SSL7 to C5 were rare

In the SSL7 blocking antibody assay, human sera were tested for their ability to block binding of immobilised SSL7 to C5. It was found that 7% of cases of SAB and 24% of healthy controls had sera which blocked binding of SSL7 to C5 [Figure 10.21B]. The potential significance of the difference in the presence of blocking antibody needs to be interpreted carefully; when the definition of a positive result was extended to include 25% of the people whose results indicated the lowest C5 binding to SSL7, there was no difference between cases with SAB and healthy controls. It is important to note that there was minimal difference between the optical density measurements of the cut-off derived from the healthy controls and the cut-off derived from the 25% lowest results.

In order to further investigate the functional activity of anti-SSL7 antibodies a second assay was developed to test the ability of patient sera to neutralise SSL7 mediated inhibition of haemolysis. The SSL7GL10 allele was used in the blocking antibody assay and the neutralisation assay, because this allele was a potent inhibitor of complement mediated haemolysis. It was found that anti-SSL7 antibody that neutralised SSL7 mediated inhibition of haemolysis was rare. There was a weak correlation between each subject's results and the results of the blocking assay. The SSL7 neutralisation assay was difficult to optimise, and the sensitivity of the assay was reduced by increasing the concentration of SSL7GL10, which reduced the amount of background haemolysis. This might have resulted in an inability of physiological concentrations of anti-SSL7 antibodies to neutralise the function of SSL7, although increasing the concentration of test sera in the assay did not alter the results.

The low prevalence of antibodies that either blocked SSL7 binding to C5 or neutralised SSL7 mediated inhibition of haemolysis is surprising, considering that high concentrations of anti-SSL7 antibody were detected in all sera tested. Laursen et al, indicated that the affinity between SSL7 and IgA and C5 is very high and neutralising antibody would need to have similar high affinity.

It is possible that anti-SSL7 IgG is directed against the IgA binding sites of SSL7, and these were not tested in the blocking and neutralisation assays. Sera from many of the cases of SAB appeared to augment SSL7 mediated inhibition of haemolysis; in two-thirds of cases the haemolysis results were less than half that of the fetal calf serum negative controls. The ability of anti-SSL7 IgG to augment SSL7 mediated inhibition of haemolysis was predicted from the crystal structure of SSL7-C5 by Laursen et al who speculated that IgG antibodies might aid SSL7 to block access of C5 convertase to C5.
The interactions between SSL7 and C5 are illustrated in Figure 10.21. C5 convertase [either C4b2a3b via activation of the classical complement pathway or C3bBb3b via activation of the alternative complement pathway] cleaves C5 producing a large fragment, C5b, which associates with C6, C7, C8 and multiple copies of C9 to form the membrane attack complex [MAC] and C5a which binds to myeloid cells resulting in their activation [Figure 6.28A]. SSL7 binds C5b and prevents formation of the membrane attack complex [Figure 6.28A]. Prevention of the formation of C5a is the most important role of SSL7. The Fc portion of IgA forms a pentameric complex with two SSL7 proteins and two intact C5 proteins which are bound in such a way as to prevent cleavage of C5 by C5 convertase thus inhibiting formation of C5a [Figure 6.28B]. IgG with high affinity for the C5 binding site of SSL7 may prevent binding of SSL7 to its target site on C5 and, thus, allow cleavage of C5 to C5a and C5b with consequent formation of MAC [Figure 6.28C]. Such antibody mediated blocking of SSL7 mediated inhibition of haemolysis was rare with the sera tested in this study. Alternatively, IgG might bind to the IgA binding site on SSL7, forming a pentameric complex similar to that centred on the Fc portion of IgA, and similarly prevent cleavage of C5 by C5 convertase, thus inhibiting the formation of MAC [Figure 6.28D]. The finding that sera from two thirds of cases with SAB augmented SSL7 mediated haemolysis may result from the formation of IgG-SSL7-C5 pentamers.

Figure 6.28 The interaction of SSL7, C5, and IgA or IgG; [A] C5 is cleaved by C5 convertase resulting in the formation of C5a and C5b. C5b interacts with other late components of the complement cascade [Fig 6.1] to form the membrane attack complex [MAC]. The role of C5b in the formation of the MAC is blocked by binding of SSL7 to C5b.
Figure 6.28 continued; [B] SSL7 forms pentamers with C5 and IgA which prevents access of C5 convertase to C5 and the formation of C5a and MAC; [C] high affinity IgG antibodies [present in a small proportion of people] bind SSL7 and prevent SSL7 mediated inhibition of C5 convertase action. C5 cleavage results in MAC formation and C5a recruitment of myeloid cells.
6.5.4 Selection pressure in ssl7 and lack of allelic variation in anti-SSL7 antibodies

Analysis of the alignment of 94 ssl7 nucleotide sequences indicated several codons under negative [purifying] selection and one site under positive selection [Chapter 6]. The sites under negative selection were all found in the OB domain at sites important for binding to IgA [N38, N68 and L109, numbered from the mature +1 residue]. In a recent report, a triple SSL7 mutant [N68T-L109A-P112A] had reduced binding to IgA. This reduction in IgA binding did not alter the mutant SSL7’s ability to inhibit haemolysis when compared with wild type SSL7, despite the essential role of IgA binding in inhibition of C5 convertase mediated cleavage of C5. Negative selection at these sites suggests that IgA binding is of critical importance to the function of SSL7 and point mutations that alter the amino acid sequence at these sites are not maintained in ssl7. It is also plausible that the amino acids at these sites are maintained under negative selection in ssl7 if IgG binding to these epitopes augments SSL7’s function.

The site under positive selection [D151] has been shown to be important for SSL7 binding to C5. A mutant at this site [D151A] reduced the ability of SSL7 to inhibit the formation of the membrane attack complex. Non-synonymous changes to codon 151 [that is, those that result in a change in the translated amino acid] have been found more frequently than would be expected if the site was not under positive selection, suggesting that frequent changes to D151 have enhanced
escape from host immunity. The presence of a wide variety of epitopes in the region where SSL7 binds to C5 could explain why concentrations of high affinity anti-SSL7 IgG are rarely sufficient to block binding of SSL7 to C5 or to neutralise the SSL7 mediated inhibition of haemolysis.

6.5.5 Anti-SSL3 antibodies were increased during SAB

Because antibody concentrations to SSL proteins were high in all test samples, the concentrations could not be correlated with the clinical or demographic features of the host.

Concentrations of serum anti-SSL3 antibody were lower in healthy controls than in cases of SAB; however, all of the healthy controls had high concentrations of anti-SSL3 antibody. This difference between cases and controls might have been the result of immunological boosting in cases with SAB, whose anti-SSL3 concentrations remained high during convalescence. The clinical significance of this finding remains uncertain, the function of SSL3 is under investigation and a functional antibody assay is not yet available to determine whether the difference in anti-SSL3 concentrations between cases and controls is associated with a difference in inhibition of SSL3’s action.

6.5.6 The limitations of radio-immunoassay for anti-SSL5 antibodies

Measurement of anti-SSL5 antibody in sera from cases and controls was abandoned due to an inability to distinguish the results of the test samples from the negative control values. This was not due to binding of SSL5 to sialic acid residues on the Cowan I SA cells used as a protein A source, but was likely to be the result of interactions between non-specific antibodies and the highly positively charged SSL5 protein. It was less likely that this was the result of interaction between SSL5 and the Cowan I cells as the negative control values were almost ten times lower than the positive control values and the Cowan cells were present in excess.

6.5.7 Ethnic variation in anti-TSST-1 and anti-SSL9 antibody concentrations

The mean concentrations of antibody against SSL9 and TSST-1 were lower in Māori and Pacific people than in people of other ethnicities; however, the mean concentrations of antibodies against SSL3, SSL7, and SSL11 did not differ between ethnic groups. The clinical relevance of the difference in anti-SSL9 antibody concentrations is unknown; all ethnic groups had high concentrations of antibody to SSL9. However, the difference between ethnic groups in antibody concentrations against TSST-1 is likely to be significant: a marked difference was also observed between Māori and Pacific cases and cases of other ethnicities in the mean neutralisation of TSST-1 mediated T cell proliferation.

The differences were independent of the age of the cases in each ethnic group, even though there was a weak positive correlation between anti-TSST-1 antibody concentrations and
increasing age. The differences observed were not due to skewing of the results caused by unusually low results from some individuals. The difference persisted even when the lowest 25% of results were removed from the analysis; and the proportion of Māori and Pacific cases of SAB whose serum samples achieved near maximal neutralisation of TSST-1 [inhibition of >95% of T cell proliferation] was lower than the proportion of cases of other ethnicities [Māori and Pacific case 34%, cases of other ethnicities 62%].

The difference in TSST-1 antibody concentrations was not related to any clinical factors, including renal failure. Any significant differences in neutralisation of TSST-1 related to indicators of deprivation were caused by confounding due to ethnic variation in those indicators; furthermore the differences were not biologically plausible. For example, Māori and Pacific people were less likely to own their homes than people of other ethnicities, and home owners had higher mean neutralisation of TSST-1 [84%] than non-home owners [72%]. However, there were no differences between home owners and non-home owners within each ethnic group. Previous studies of the role of TSST-1 in SAB have not found any relationship between concentrations of serum antibody to TSST-1 and susceptibility to SAB. In this light it is worth remembering that patients with toxic shock syndrome due to SA infection, who lack antibody to TSST-1, very rarely have SAB, but have localised mucosal or wound infection. Thus, lack of anti-TSST-1 antibodies that neutralise TSST-1 mediated T cell proliferation does not appear to predispose to SAB.

The reasons for the difference in TSST-1 neutralisation between ethnic groups could not be determined in the current study, but is an important area for future research. Although anti-TSST-1 antibody concentrations were not related to outcome among cases of SAB in the current study, a recent report has suggested that tsst positive strains are more commonly associated with SAB. Kansal et al, determined that anti-TSST-1 IgG1 antibody was more closely correlated with neutralisation than other immunoglobulin isotypes; and they speculated that host genetic factors control the type of antibodies that are formed in response to an infection. It is not known whether there are differences in the IgG subclasses that respond to SA infection between different ethnic groups. The development of the T cell repertoire is not entirely random; however, the near random nature of variable immunoglobulin chain reassortment in the thymus is unlikely to account for the difference observed between ethnic groups.

6.5.8 Summary

This study has shown that antibodies against SSL proteins were ubiquitous, but antibodies against SSL7 did not impair SSL7’s interaction with complement factor C5. The finding of ethnic variation in antibodies against TSST-1 was surprising, but this difference was also seen in anti-SSL9 antibody concentrations. Further research is required to understand the reasons for this variation, which may have significant implications for the prevalence and prevention of a number of infectious diseases in our region.
Chapter 7: Conclusion and future directions

The results presented in this thesis indicate that almost 800,000 New Zealanders have SA nasal carriage at any one time; approximately 75,000 New Zealanders develop skin and soft tissue infection each year; and almost 800 develop *S. aureus* bacteraemia [SAB] each year. Māori and Pacific people have higher rates of SAB than people of other ethnicities. The reasons for ethnic variation in the incidence of SAB need to be examined; even though Māori and Pacific cases of SAB are younger than cases of other ethnicities, their mortality rate remains very high [15%].

There are several potential reasons for this striking difference in the incidence of SAB between ethnic groups.

### 7.1 The prevalence of nasal colonisation might be higher in Māori and Pacific people

It is likely that SA carriage precedes the development of SAB; for example, hospital-acquired SAB is more common in nasal carriers of SA \(7^1\). If the model of SA disease arising from nasal carriage, which occasionally progresses to skin and soft tissue infection and/or invasive disease, holds true, then the proportion of each ethnic group that has SAB or skin and soft tissue infection is expected to be similar to the proportion of each ethnic groups that has SA nasal carriage. In Auckland the prevalence of SA nasal carriage did not differ between ethnic groups. Therefore, the observed ethnic variation in the incidence of SAB is not caused by ethnic variation in the prevalence of SA nasal carriage. The number of Pacific people who have wound swab cultures performed is in proportion with the higher prevalence of SAB in Pacific people; thus it is possible that colonisation with SA at a site other than the anterior nares is often the precursor of skin and soft tissue infection and SAB, particularly in Pacific people.

### 7.2 Colonisation with particularly virulent strains of SA might be more frequent in Māori and Pacific people

Higher rates SAB in Māori and Pacific people were not related to higher rates of SA colonisation. However, the SA strains that colonise Māori and Pacific people might be more virulent and more likely to cause SAB. Genetic analysis of the SA population in Auckland, showed broad similarity between the SA populations that caused infection in Māori, Pacific people and people of other ethnicities. However, one lineage of SA, CC121, caused infection only in Māori and Pacific people; this lineage is common in other parts of the world, but the identification of a novel CC121 variant, which has its origins in the Pacific, suggests that some CC121 strains might have
adapted to Māori and Pacific people. Furthermore, the MRSA strains that cause skin and soft tissue infection in Māori and Pacific people differ from the MRSA strains that cause infections in people of other ethnicities. MRSA is a rare cause of nasal carriage [<1%], but a much more common cause of skin and soft tissue infection [8%] and SAB [7%]. A larger study of nasal SA carriers might show higher rates of MRSA colonisation in Māori and Pacific people, but it is also possible that nasal carriage is not the main precursor to development of more serious disease.

The high rates of travel and migration between Auckland and Pacific Island nations, coupled with the differences between MRSA strains found in Samoa and New Zealand, raised the expectation that further important differences would be identified between SA populations in Auckland and Pacific Island nations. However, the SA populations of Samoa, Tonga and Fiji do not differ significantly from the SA population of Auckland. Surprisingly, the SA populations obtained in small Pacific Island nations contained the same degree of genetic diversity as SA populations from other countries around the world; furthermore the SA populations from the Pacific nations contained the same core genotypes of SA found in SA populations from other countries around the world. Given the rapid rate that SA evolves, this finding suggests that SA populations have recently become rapidly mixed.

7.3 The use of invasive devices, which increase the risk of invasion by colonising organisms, might be higher in Māori and Pacific people

Māori and Pacific people are more likely to have end-stage renal failure and diabetes than people of other ethnicities, and these confounding factors contribute significantly to the higher prevalence of SAB in Māori and Pacific people. The increased prevalence of healthcare-associated SAB among Pacific people is due to infection of central venous catheters used for haemodialysis, which are known to be more commonly used for haemodialysis among Māori and Pacific people. However, infection of haemodialysis catheters does not adequately explain the higher prevalence of SAB in Māori people nor the higher prevalence of community-onset SAB in Māori and Pacific people.

7.4 Innate or adaptive immune responses against SA might be less protective against invasive disease

The simplest explanation for the development of invasive SA disease involves exposure to a strain of SA, to which the person lacks effective immunity. Māori and Pacific people have lower concentrations of antibody against SSL9 and TSST-1. In the case of TSST-1, low antibody concentrations correlated with a lower degree of neutralisation of TSST-1. Increased susceptibility to SAB in Māori and Pacific people is not directly caused by reduced anti-SSL9 or anti-TSST-1 antibody concentrations; the antibody concentrations were not significantly different between cases of SAB and healthy controls, and many of the episodes of SAB would have been
caused by tsst negative strains of SA. However, this finding indicates the need to measure antibody concentrations against other SA virulence factors potentially associated with increased susceptibility to SAB amongst Māori and Pacific people.

7.5 Māori and Pacific people might develop mild SA disease that might be neglected or undertreated and progress to SAB

In Auckland, Pacific people are more likely to have a wound swab culture performed compared with Māori and people of other ethnicities. Thus, it is possible that Pacific people are more likely to develop mild skin and soft tissue infections that progress to SAB. It is equally possible that healthcare providers are aware that some MRSA strains are more commonly isolated from wounds of Pacific people and are more likely to request a wound swab culture in addition to providing antimicrobial treatment. Regardless, this finding suggests that Pacific people have adequate access to healthcare. Māori and Pacific cases of SAB were not less likely to be registered with a GP, nor were they less likely to see their GP in the month prior to their episode of SAB. Opportunities to prevent SAB might be infrequent; only a few cases of SAB reported any symptoms of SA infection prior to the onset of the illness associated with their episode of SAB. All of these cases had skin and soft tissue infections that progressed rapidly, resulting in admission to hospital with SAB.

7.6 Ethnic variation in rates of SAB might be caused by increased susceptibility and exposure to SA due to socioeconomic deprivation

This thesis found that cases of SAB are more socially deprived than the total population and community-onset cases of SAB, which were not associated with renal failure or diabetes, were more likely to be Māori or Pacific people that live in crowded households within areas with the highest NZDep2006 scores. It is difficult to conclusively state that deprivation causes ethnic variation in the prevalence of SAB, because indicators of social deprivation also vary between ethnic groups in society as a whole. Any illness that is more common in Māori and Pacific people will also be more common in more deprived people and any illness that is associated with deprivation will appear to be more common in Māori and Pacific people.

Social deprivation is said to be an important risk factor for the development of diseases caused by Streptococcus pyogenes and Neisseria meningitidis. A retrospective audit of rheumatic fever cases in the Waikato region found that 62/77 [81%] cases were Māori and 57/73 [78%] of these cases lived in areas with the highest 3 deciles of NZDep2001 scores. During a recent epidemic, the high prevalence of meningococcal disease in Māori and Pacific children was largely due to high rates of household crowding. Studies that have investigated the role of social deprivation in the development of rheumatic fever, meningococcal disease, and, in this thesis,
SAB and SA nasal carriage have great difficulty distinguishing between the effects of ethnicity and deprivation.

It is likely that deprivation leads to higher rates of infectious disease via increased pathogen exposure. SA carriage, and presumably exposure, peaks in early childhood, yet the incidence of SAB in children under the age of 16 years [16.9/100,000/year] was less than the incidence in adults over 16 years [41/100,000/year] during the same time period in the same cities in New Zealand. 

7.7 Questions for future research

The studies presented in this thesis have highlighted the importance of SA as a pathogen in New Zealand, and have shown that ethnic variation in the prevalence of SAB does not have a straightforward explanation. Higher rates of SAB are not caused by higher rates of nasal colonisation. Infection of haemodialysis catheters contributes to the high rate of SAB in Pacific people, but is less of an explanation for the increased prevalence of SAB in Māori people. Deprivation is an important factor in the development of SAB, but the differences identified between ethnic groups reflect differences in society as a whole. Immune responses to SA differ between ethnic groups but this is not due to infection with different strains of SA; there are no substantial differences between the SA that infect Māori and Pacific people or people of other ethnicities. There are five important questions for future research that have arisen from the analysis presented in this thesis:

1. Are practices regarding infection control in haemodialysis units, the management of haemodialysis vascular access and the use of SA decolonisation optimal?

Reduction in the prevalence of central venous catheter use among people with end-stage renal failure would be expected to immediately reduce the prevalence SAB and prevent deaths. Selective decolonisation of SA carriers among haemodialysis recipients might reduce the incidence of SAB. Recently, Bode et al showed that decolonisation of nasal carriers of SA with topical intranasal mupirocin and chlorhexidine body wash was effective in reducing hospital-acquired SA infections in patients admitted to medical and surgical wards, but their large study lacked the power to show a reduction in SAB. Haemodialysis recipients might differ from general hospital inpatients, the presence of long term central venous lines, high burden of SA infection and frequent healthcare exposure might reduce the efficacy of SA decolonisation. There are no randomised controlled trials that show a reduction in SAB among haemodialysis recipients following SA decolonisation, and Laupland and Conly concluded that decolonisation using intranasal mupirocin was not a recommended evidence-based practice. However, two studies have shown that the prevalence of SAB among haemodialysis recipients reduced following topical intranasal mupirocin when compared to the historical prevalence of SAB. Thus, a study to investigate the utility of intranasal mupirocin and chlorhexidine body wash among
haemodialysis recipients in New Zealand is a high priority, and might show a reduction in the prevalence of SAB in these high risk people.

2. Are there substantial differences between SA strains that cause nasal colonisation and SA strains that cause SAB?

Nasal carriage strains differ from invasive disease strains; which possess mobile genetic elements that aid virulence \(^95,113\). Sivaraman et al compared the genomes of two related SA isolates that caused nasal colonisation and invasive disease, but no studies have been reported that examine the genomes of multiple related SA isolates to compare nasal carriage isolates with invasive disease isolates. It is quite likely that the mortality of SAB has not changed considerably over recent decades despite advances in healthcare. Further understanding of the bacterial factors that are critical for the development of SA invasion and immune evasion could lead to therapies to target those factors to prevent SA disease. The SA strains collected for this thesis, which included nasal carriage isolates and SAB isolates obtained from the same population over the same time period, provide an ideal starting point to examine this question in detail. Genomic comparison of ST1 strains that cause nasal carriage and SAB could identify bacterial factors that are consistently associated with nasal carriage and disease.

3. How does socioeconomic deprivation contribute to the development of infectious diseases in New Zealand?

SA infection is an excellent disease model to study, it is common and straightforward to diagnose. It has advantages over rheumatic fever, which is uncommon in non-Māori and non-Pacific children; severe \(S. pyogenes\) sepsis is also uncommon; meningococcal disease rates are highly correlated with young age and this disease is now uncommon following resolution of the recent NZ meningococcal epidemic. The exact mechanism by which social deprivation causes infectious diseases in New Zealand is not entirely understood. In order to inform social and health policy this is a crucial area for future research. A large prospective case-control study of SA skin and soft tissue infection could compare a large number of factors, including indicators of social deprivation, between Māori people with and without disease, Pacific people with and without disease etc. The study would also provide an opportunity to measure the frequency and rate of SA transmission, which remains largely unknown.

4. What are the origins of contemporary SA populations?

Figure 7.1 shows the STs of SA in the MLST database grouped into clonal complexes by goeBURST 1.2.1. On a global scale, the largest clonal complexes that contain the highest number of STs are either the oldest or the most successful. CC30 contains several smaller groups of STs that could be considered to be CCs in their own right: CC34 and CC39. In contrast CC121 contains fewer STs and is either younger or less successful on a global scale.
The origins of CC30 could be examined by comparison of CC30 strains from around the world [including the CC30 strains analysed in this thesis]; this study would provide important clues regarding the origins of SA as a human pathogen. This study would require a tool with greater resolution than MLST and recent comparisons of SA genomes have shown that next generation sequencing and computational analysis now provides the opportunity to perform such a study\textsuperscript{188, 190, 198}. Examination of a global sample of methicillin-susceptible CC121 strains is expected to provide insight into the degree of adaptation of CC121 strains to Māori and Pacific people in comparison with the CC121 that cause infection in other ethnic groups elsewhere in the world. If CC121 is younger than CC30 then comparison of CC121 and CC30 strains will provide further understanding of the evolutionary processes that give rise to a new and successful SA strain, that, unlike ST239 does not appear to have been driven by antimicrobial selection pressure\textsuperscript{188}.  

\begin{figure}[h]  
\centering  
\includegraphics[width=\textwidth]{fig7.1.png}  
\caption{Global SA sequence types grouped into clonal complexes (CC) that share at least 5 of 7 identical MLST loci by goeBURST 1.2.1.}  
\end{figure}
5. What is the mechanism for ethnic variation in antibody response to TSST-1 and SSL9?

There is no simple explanation for the variation in immune response to TSST-1 and SSL9 between ethnic groups, but research regarding this finding could have important implications for ethnic variation in the prevalence of infectious diseases caused by other pathogens, such as *S. pyogenes*, *S. pneumoniae* and *N. meningitidis*. If Māori and Pacific people are more susceptible to SAB due to variation in immune response to SA antigens, then study of immune responses to other SA antigens in people living in Auckland might be able to identify the main antigen[s] responsible for the development of SAB. This finding would be an important step in the identification of vaccine targets and vaccine development is the ‘holy grail’ of SA research: opportunities to prevent SAB by other means are limited and the mortality of SAB remains high, despite modern healthcare.
Appendix A: Example of population genetic analysis using ssl11 gene fragment

This Appendix contains a brief analysis of a short fragment of the ssl11 gene for 15 SA strains. This gene fragment contains more polymorphism than the MLST loci. This example provides the reader with insight regarding the interpretation of population genetic analysis complementary to the results presented in Chapter 5.

A.1 Example of calculation of the genetic distance between nucleotide sequences

Figure A.1 shows a short nucleotide alignment from a segment of the open reading frame of ssl11 for 15 different isolates of SA.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Nucleotide Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>2</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>3</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>4</td>
<td>ATG TCT TCT TGA A T</td>
</tr>
<tr>
<td>5</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>6</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>7</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>8</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>9</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>10</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>11</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>12</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>13</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
<tr>
<td>14</td>
<td>ATG TCT TCT TGA A G</td>
</tr>
<tr>
<td>15</td>
<td>ATG TCT TCT TGA A C</td>
</tr>
</tbody>
</table>

Figure A.1. Nucleotide alignment of 15 nucleotide segments from the ssl11 gene. Polymorphic sites are highlighted in boldface text.

Isolates 1 and 2 differ at 12/63 (PWD = 0.19) sites in the nucleotide sequence but isolates 2 and 3 differ at 6/63 (PWD = 0.10) sites. A matrix of genetic distances is required for a number of analyses; an example from the first five sequences in Figure A.1 is shown in Figure A.2.
<table>
<thead>
<tr>
<th>isolate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.19</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.19</td>
<td>0.10</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.22</td>
<td>0.25</td>
<td>0.22</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.24</td>
<td>0.10</td>
<td>0.11</td>
<td>0.22</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure A.2. Matrix of genetic distance for the first five nucleotide sequences from a segment of the *ssl11* gene from 15 isolates of SA.

A.2 Example rarefaction curve

If the *ssl11* sequences shown in the matrix in Figure A.2 were grouped into operational taxonomic units (OTU) that contained isolates related by a PWD of less than 0.15 [OTU < 0.15] then isolate 1 would not form an OTU < 0.15 with any of the other 5 strains, and thus, would form an OTU < 0.15 on its own. Isolates 2, 3 and 5 would form another OTU < 0.15 as would isolate 4.

A rarefaction curve is derived from drawing random samples from the population without replacement and calculating the number of different OTUs present after each sample has been drawn. The procedure is repeated 1,000 times, and an example of a rarefaction curve for the *ssl11* gene segments is shown in Figure A.3. Figure A.3A shows the rarefaction curve for OTU unique as well as 95% confidence intervals from 1000 random iterations. The curve continued to rise as samples were randomly drawn from the 15 isolates; and, by extrapolation, there are a number of OTU unique that haven’t been sampled. Figure A.3B shows the rarefaction curve for OTU < 0.15; the rarefaction curve for OTU < 0.15 indicates that all five OTU < 0.15s were identified after an average of ten random draws from the 15 nucleotide sequences; the rarefaction curve became saturated at this point. It was unlikely that the diversity or richness of the sample [defined by strains that differ by a PWD of less than 0.15] would be improved by additional sampling.
Figure A.3. The rarefaction curves for [A] OTU_{unique} with 95%CI and [B] OTU_{0.15} for nucleotide segments of the ssl11 gene from 15 isolates of SA.

A.4 Example of Efron’s calculation and Simpson’s index of diversity

The reward of additional sampling estimated by Efron’s calculation was checked using the Shen and Solow formulae; all implemented in MOTHUR\textsuperscript{146}.

For example, if an additional 15 isolates were sampled, only one OTU_{unique} would have been identified for the ssl11 gene segments. Doubling the sample size would not have added any OTU_{0.15}s.
= 5 - 3 + 0 - 1 \\
= 1.

Simpson’s index was calculated to estimate the diversity ssr1 gene segments from 15 isolates of SA, which contained five OTU_{15}s. Three of these OTUs had two members, one had three members and one had 6 members.

\[
\text{Diversity} = \frac{3[2][2-1] + 1[3][3-1] + 1[6][6-5]}{15[15-1]} \\
= \frac{42}{210} \\
= 0.2
\]

**A.5 Example of estimation of recombination and mutation: LDHAT**

16 segregating sites were identified in the ssr1 nucleotide segments shown in Figure A.1. In total there were 120 pairs of segregating sites, which were tested for their likelihood using a range of possible values of \( \rho \) from 0 to 100. The calculated value of \( \theta \) [\( \theta=4.921/63 \) nucleotides = 0.078/nucleotide] was used throughout. The maximum composite likelihood score was -2651 and this occurred when \( \rho \) was 18 [Figure A.4]. \( \tau/m \) was calculated from \( \rho \) divided by \( \theta \) and represents the relative rate which a nucleotide may vary by recombination or mutation. For the example this equals 3.66 [\( \rho=18/63 \) nucleotides divided by \( \theta=4.921/63 \) nucleotides].

![Figure A.4](image.png)

*Figure A.4. The composite likelihood curve for different values of \( \rho \) [rho] for 16 segregating sites from segments of the ssr11 gene from 15 isolates of SA. The maximum likelihood occurred when \( \rho=18 \).*
A.6 Example of detecting recombination

The LDHAT likelihood permutation test was used to detect the presence of recombination. The power of the test is improved if only segregating sites with 2 alleles, where the minor allele has a frequency of >10%, are analysed. 16 segregating sites were identified in the ssl11 gene segments. The first of these was at position 4 where 13% [2/15] of the isolates contained a thymine at that position and the remainder adenine. There were also two alleles evident at position 3, but only 1/15 [7%] was polymorphic, which was below the 10% threshold. Likewise, position 27 is not included because there were three alleles at this site: guanine, thymine and cytosine. In the previous example, the maximum composite likelihood for the value of $\rho$ for the ssl11 gene segments was $-2651$. When the segregating sites were shuffled 1000 times a likelihood value equal to or greater than this was found 31 times [i.e. $p=0.031$], indicating that shuffling reduced the likelihood measurement because some sites were linked.

The PHI test was not able to be implemented as the nucleotide sequences were too short and contained too few segregating sites.

A.7 Example of finding recombination breakpoints using GARD

GARD identified 24 potential breakpoints for the ssl11 gene segments. A large number of models were tested and there was strong evidence for one breakpoint, at nucleotide site 37. However, the likelihood values of the NJ trees constructed from site 1 to 36 and from site 37 to 63 were not greater than the likelihood of the NJ tree constructed from the total alignment, so this breakpoint was rejected [Figure A.5].
Figure A.5. The phylogeny [NJ tree] reconstructed for [A] the total nucleotide alignment and [B] from nucleotide sites 1-36 for ssl11 gene segments from 15 isolates of SA. The likelihood value of each tree was not significantly different.

A.8 Example of estimating selection pressure for nucleotide sequences

The ssl11 gene segments shown in Figure A.1 each contain 63 nucleotides encoding 21 amino acids. The 25 polymorphic nucleotides result in 15 variable codons and 6 non-variable codons. In total, the alignment contained 23 synonymous and 24 non-synonymous substitutions, which were used to determine dS and dN at each site. Figure A.6 shows the difference between dN and dS at each codon, normalised under the assumption that all substitutions in the alignment are under a binomial distribution. Codons without bars did not contain any polymorphism; there were four non-synonymous substitutions at codon 6, but this value did not reach the p value cutoff of 0.1, and this site was not considered to be under positive selection. Negative selection was identified
at three codons with negative values of dN-dS [9, 11 and 12] which exceeded the significance value. The average values across the alignment gave a dN/dS ratio of 0.22.

Figure A.6. The difference between the normalised rate of non-synonymous and synonymous nucleotide substitution (dN - dS) for each ssl11 gene fragment codon; the light grey bars indicate codons where significant selection was detected.

A.9 Example of the parsimony test and LibShuff

As an example, imagine that the ssl11 gene fragments shown in Figure A.1 were isolated from SA obtained from two groups of people wearing blue or red shirts [a nonsense scenario invented to aid the example]. The two groups are indicated by the NJ tree shown in Figure A.7. To test the null hypothesis that there was no difference in the ssl11 gene segments isolated from people wearing blue or red shirts the Parsimony test was used. The parsimony score between the two populations was 2. The tree tips were then shuffled randomly 1000 times and a parsimony score of 2 or less was found on only four occasions. This value [p=0.004] indicated that the two populations do occupy different topologies on the NJ tree and the populations are different.
Figure A.7. The phylogeny of *ssl11* gene segments from 15 isolates of SA isolated from 8 people wearing blue shirts [blue] and 7 people wearing red shirts [red].

The LibShuff method was also used to compare the *ssl11* gene segments from SA isolated from people wearing blue shirts and from people wearing red shirts [as indicated in Figure A.7]. First a curve of coverage versus distance was drawn for each population [Figure A.8]. The difference between points on each curve was squared and then summed to determine $\Delta C$. This score was 0.21 for the difference between curves of SA isolates obtained from people wearing blue shirts and red shirts. Next, the two groups were shuffled 10,000 times and $\Delta C$ was calculated for each shuffle. If the nucleotide sequences from SA obtained from people wearing blue shirts were significantly different, a random shuffle should reduce the difference between the curves of coverage versus distance and provide lower values of $\Delta C$ most of the time. However, a value of $\Delta C$ equal to or greater than 0.21 was found after more than 4900 shuffles, indicating that the genetic distance between nucleotide sequences obtained from people with blue shirts was not different to the genetic distance between sequences obtained from people wearing red shirts.
Figure A.8. Coverage versus distance for ssl11 gene segments obtained from isolates of SA obtained from people wearing blue shirts [blue] and red shirts [red].

A.10 Example of 2 dimensional multi-dimensional scaling

The two dimension MDS map of ssl11 gene segments obtained from SA isolated from people wearing red shirts and blue shirts is shown in Figure A.9. The strains with identical sequences are plotted on top of each other and the stress score was low, 0.06 [Figure A.9A]. This value was obtained in 26 of the 50 iterations. Figure A.9B shows the scatter plot of the distance of the gene sequences on the MDS plot versus the genetic distance of the isolates. The line of best fit is shown in Figure A.9B, and the stress value reflects the distance of the data points from the line of best fit. This MDS plot indicated a difference between the ssl11 gene segments from SA isolated from people wearing blue shirts and people wearing red shirts.
Figure A.9. [A] The 2-D MDS plot of ssl11 gene segments from SA isolated from people wearing red and blue shirts; and [B] the scatter plot from the final MDS map of MDS distance versus genetic distance; the line of best fit [grey line] is shown for data points [blue circles].

A.11 Example of ANOSIM and PERMANOVA

When ANOSIM was performed to test the null hypothesis of no difference between ssl11 genes from people wearing blue shirts and people wearing red shirts the R value was high, 0.681. This value indicated that the null hypothesis was unlikely. All possible permutations were performed
and R did not exceed 0.6 for any of these permutations, thus the null hypothesis was rejected \(p<0.05\). Figure A.10 shows the frequency distribution of R values for the ANOSIM permutation test; the true R value is indicated by the vertical line.

Figure A.10. The frequency distribution of R values for ANOSIM permutation testing of differences in ssl11 gene segments from SA isolated from people wearing blue shirts and red shirts. The true value of R [0.681] is indicated by the vertical line.

PERMANOVA was used to test whether there was a relationship among ssl11 gene segments from SA and the shirt colour and dominant hand of the human host. The MDS plot with the additional factor of dominant hand is shown in Figure A.11A. To test the null hypothesis that there was no difference in ssl11 gene segments from left or right-handed people the genetic distance matrix was split into two partitions and the sum of squared values among groups was calculated [342.87] which was the same as the mean squared value as there were only two groups for this variable [among left and right-handed people; \([a-1]=1\)]. The mean squared value within the groups was 42.154, which was \(SS_{\text{residual}} [548]\) divided by the residual degrees of freedom [\([N-a]=13;\) there were 14 comparisons between the 15 isolates]. Thus, pseudo-\(F\) was 8.13 [342.87/42.154]. 10,000 permutations were then performed and this value of pseudo-\(F\) was only reached or exceeded four times \(p=0.0004\), there was a significant difference in ssl11 gene fragments isolated from left or right handed people. In the previous example, ANOSIM found a significant difference in ssl11 gene fragments isolated from people wearing different coloured shirts. PERMANOVA was used to explore the interaction among these two variables. After shirt colour was taken into account the handedness of the host was no longer significant [pseudo-\(F=1.61, 1\) d.f, \(p=0.2054\)] nor was the interaction among shirt colour and handedness [pseudo-\(F=1.0588, 1\) d.f, \(p=0.3697\)] [Figure A.11B].
source & degrees of freedom & Sum of squared distance & Mean sum of squared distance & Pseudo-F & p value \\
--- & --- & --- & --- & --- & --- \\
shirt colour & 1 & 372.99 & 372.99 & 9.8447 & 0.0001 \\
dominant hand & 1 & 60.999 & 60.999 & 1.61 & 0.2054 \\
shirt colour x & 1 & 40.114 & 40.114 & 1.0588 & 0.3697 \\
dominant hand & & & & & \\
residual & 11 & 416.76 & 37.887 & & \\
total & 14 & 890.87 & & & \\

**A.12 Example of phylogeny reconstruction using maximum likelihood: TREEPUZZLE**

The ML tree for the ssl11 gene segments described above is shown in Figure A.12A. This tree concurred with the previous analyses that showed a difference in the SA isolated from people wearing different coloured shirts. All of the 1365 possible quartets for these sequences were examined, and 17.4% of them were unresolved [Figure A.12B]. A further 7% [2.2% + 2.2% + 2.6%] were partially resolved, i.e. two possible topologies were equally likely. Thus, this tree needs to be interpreted with caution; it was possible that some of the model assumptions were incorrect or that recombination, deletion, or rearrangement had altered the ability to reconstruct the phylogeny based on a model of random point mutation.
A.13 Example of ancestral population using STRUCTURE

For the ss/11 gene segments 26 polymorphic sites were examined by STRUCTURE. \( \lambda \) was found to be 2.0788 and 2.0912 for two iterations of K=1 and \( \lambda \) was set at the mean value [2.085] for the remainder of the analysis. Three iterations of different values of K from 1 to 12 were performed. The likelihood values for each iteration and the value of K is shown on Figure A.13A. The likelihood of K=2 was greater than K=1, but then declined at higher values of K. Thus the genetic data identified two ancestral populations. A more complete run of three iterations for K=2 was performed, for 500,000 repetitions and 50,000 burn-in repetitions. These three runs were pooled using CLUMPP and then displayed using DISTRUCT [Figure A.13B, C and D]. Figure 5.17B shows the proportion of inheritance from two ancestral populations for two groups of ss/11 gene fragments isolated from people wearing red or blue shirts. Figure 5.17C shows the inheritance for each individual isolate and then Figure A.13D shows the individuals grouped according to their
shirt colour. These plots show that *ssl11* gene segments from people with blue or red shirts derives their ancestry from two sources, and that three isolates [3, 9 and 11] contain a substantial amount of genetic material derived from both ancestral sources. This type of inference was not possible using the methods of comparison described previously.

Figure A.13. [A] The likelihood, Ln[L], of the model for different numbers of ancestral populations [K]; [B] the proportion of membership of ancestral populations derived by STRUCTURE of *ssl11* gene segments isolates from people wearing blue or red shirts; [C] the individual membership of two ancestral populations for 15 isolates; and [D] the individual membership sorted by shirt colour.
Appendix B: Forms used for data collection

B.1 Consent form

School of Medical Sciences
Department of Molecular Medicine & Pathology

CONSENT FORM

Title of project: The Auckland Staphylococcal Study

Principal Investigator: Dr. Stephen Ritchie
Dept. of Molecular Medicine & Pathology, School of Medicine
Tel.: (9) 367-0000 ext 22968

Head of Department: Prof. Peter Browett
Dept. of Molecular Medicine & Pathology, School of Medicine
Tel.: (9) 373-7599 Ext. 86218

REQUEST FOR INTERPRETER

<table>
<thead>
<tr>
<th>Language</th>
<th>I wish to have an interpreter</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Māori</td>
<td>E hiaha ana ahau ki tetahi kaiwhakamaori/ kaiwhaka pakeha</td>
<td>Ae</td>
<td>Kao</td>
</tr>
<tr>
<td>Samoan</td>
<td>Ou te mana’o ia i ai se fa’amatala upu</td>
<td>Ioe</td>
<td>Leai</td>
</tr>
<tr>
<td>Tongan</td>
<td>Oku ou fiema’u ha fakatonulea</td>
<td>Io</td>
<td>Ikai</td>
</tr>
<tr>
<td>Cook Island</td>
<td>Ka inangaro au i tetai tangata uri reo</td>
<td>Ae</td>
<td>Kare</td>
</tr>
<tr>
<td>Niuean</td>
<td>Fia manako au ke faka’aoaga e taha tagata fakahokohoko kupu</td>
<td>E</td>
<td>Nakai</td>
</tr>
</tbody>
</table>

I have read and I understand the information sheet dated Feb 2007 for volunteers taking part in the study designed to investigate the role of SSL toxins in severe staphylococcal disease and the types of *Staphylococcus aureus* bacteria in Auckland. I have had the opportunity to discuss this study. I am satisfied with the answers I have been given. I have had the opportunity to use whanau support or a friend to help me ask questions and understand the study.
I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time. This will not affect my future healthcare in any way. I understand that my participation in this study is confidential and that no material, which could identify me will be used in any reports on this study. I have had time to consider whether to take part. I know who to contact if I have any questions about the study.

Do you agree to providing a blood sample(s) if asked to do so? Y  N

Do you agree to have a swab of your nose or skin performed if asked to do so? Y  N

Are you interested in the outcome of this research project? Y  N

Please send a short report to the following address:
______________________________________________________________________
______________________________________________________________________

I agree that my blood sample may be used for a future, related study, approved by an accredited NZ Ethics Committee. Y  N

I ___________________________ (full name) hereby consent to take part in this study.

Signed:

Date:

Project explained to me by:

Signature of investigator:

Date:

If you are 15 years or younger please show this sheet to your parents. Should you decide to take part in the project your parents need to sign this Consent Form.

I agree that ________________________ who is under my guardianship may participate in this research.

Signed:

Name:
(please print clearly)

Date:
PARTICIPANT INFORMATION SHEET

<table>
<thead>
<tr>
<th>Title of project:</th>
<th>The Auckland Staphylococcal Study</th>
</tr>
</thead>
</table>
| Principal Investigator: | Dr Stephen Ritchie  
Dept. of Molecular Medicine and Pathology, School of Medicine.  
Tel.: |
| Head of Department: | Prof. Peter Browett  
Dept. of Molecular Medicine and Pathology, School of Medicine.  
Tel.: (9) 373-7599 Ext. 86281 |
| Time course of study | 3 Years, until December 2009 |

You are being invited to take part in a research study. Being in a research study is voluntary, and you need to decide whether or not you want to participate. Please read the following information carefully and ask the doctor any questions you may have. You may discuss this with a friend or your GP, or seek family or whanau support, to help you understand the benefits of this study and any other explanation you may require. An interpreter can be provided if needed. Please take your time to make your decision. Your participation is entirely voluntary (your choice).

**Purpose of this project**

There are three groups of people who have been asked to participate in this study.

Group 1: You have been asked to take part in this study because you have bloodstream infection with *Staphylococcus aureus*.

Group 2: Your whanau/family member/household member has been found to have bloodstream infection with *Staphylococcus aureus*.

Group 3: You are a healthy member of the population and we would like to test whether you have *Staphylococcus aureus* in your nose and, if so, whether it is similar to the types that causes illness in other people.

*Staphylococcus aureus* is a bacterium and is a very common cause of illness in New Zealand. It lives in the nose of 20-30% of healthy members of the population; these people are “carriers” and very rarely get sick. There are many different types of this bacterium and we would like to determine which ones cause illness in Auckland. Also, it produces a number of toxins (SSLs) and we suspect that these contribute to the symptoms of illness caused by this bacterium. We would like to study the blood samples.
from participants to see whether people with this infection develop immune responses (antibodies) to SSLs.

**Your contribution to this project**

Group 1: You are in hospital and have been diagnosed with bloodstream infection. We would like to test your blood for the presence of SSLs and to see whether you from antibodies to these over time. We only need small blood samples (1-2 ml) and we can use some of the blood that has already been taken during the period of your hospital stay and is now stored in the laboratory (these samples are usually thrown away after 7 days). We would also like to take a sample of blood at 7, 14, 21 and 28 days after the diagnosis. If you leave hospital, some of these blood samples will be additional to those required by your doctors. Dr Ritchie will talk to you about the easiest way to collect blood samples from you after you leave hospital. These blood samples will be taken by Dr Ritchie at your home, or taken at your local community testing laboratory. We would also like to ask you some questions about your illness and about your household. This interview would take 10 minutes. You can take part in this study even if you don’t wish to provide blood samples.

Group 2: You are the whanau/family member or household member of someone who has bloodstream infection. We would like to take a swab of your nose to test whether you are a carrier of *Staphylococcus aureus*. The swab is taken from the front part of the nose and should not cause any discomfort. If you are 15 years or older you might also be asked to provide one blood sample to test whether you have antibodies to this bacterium. You can take part in this study even if you don’t wish to provide a blood sample.

Group 3: You are a healthy member of the population; we would like to take a swab of your nose to test whether you are a carrier of *Staphylococcus aureus*. The swab is taken from the front part of the nose and should not cause any discomfort. We would also like to ask you some questions about your health and about your household. This interview would take 10 minutes. You will not be asked to provide a blood sample.

**Benefits, risks and safety**

This study involves testing the bacteria we have obtained from you to see what type it is, and comparing it to others. The bacteria will be taken from the laboratory (if you have bloodstream infection) or will be taken from a swab of your nose. The swab is taken from the front part of the nose and is not painful or uncomfortable.

Additionally we intend to test the serum (the liquid that bathes the blood cells) of some participants for the presence of SSLs and antibodies to these toxins. Whenever possible we will use blood samples that have already been taken by your doctors. Some participants may be asked to undergo a further blood test(s). The serum will be stored until it is analysed. There will be no genetic tests performed and only serum (i.e. no genetic material) will be stored.

There are no payments or rewards for participation in this study, but we hope to answer some important questions about an infection that is very serious and common in Auckland.

If you are in hospital, participating in this project will not alter your treatment at any time. There are, therefore, no additional health risks. Likewise, if you do not wish to participate in this study, there will be no change in your treatment. You are able to withdraw from
this study at any time, without having to give a reason, and all of the samples you have provided will be removed from the study and destroyed. Withdrawal from the study will not affect your future healthcare in any way.

The study will be carried out at our research laboratories at Auckland Medical School over the next 3 years. At the end of this period, all data will be evaluated and might be sent to a scientific journal for publication. The study will also be written up in the form of a doctoral thesis.

If you are interested in the outcome of this study, please provide your address on the Consent Form and a short report will be sent to you when the data have been evaluated; this will not be possible until the analysis has been performed and may take up to 3 years.

There will be no reference to you as an individual. No material, which could personally identify you will be used in any reports on this study. We do not need your name or any information about you other than what is listed below to help us better understand the results. You do not have to answer all of the questions and you may stop the interview at any time. The information we wish to obtain includes the following:

Your age, your gender, your home address, ethnic background and place of birth, prior/other medical illnesses such as diabetes and eczema, cigarette and alcohol usage, any medication you might have received, household members’ ages and genders, household information such as total household income and car ownership.

**Storage of samples for future study**

If you have contributed a blood sample to this study you will be asked whether you agree to have your samples stored for future use. Future use would only be considered to study another aspect of staphylococcal disease, for example, if a new toxin is discovered then your sample could be tested. A future study could only be performed if an ethics committee has provided approval. The specimens will be stored indefinitely and will only be identified by a numeric system and will not be labelled with any personal information. No cells or genetic material will be stored. If you agree to have your sample(s) stored then your samples will be destroyed after they have been analysed for the current study. You can take part in the study even if you don’t consent to storage of your samples for future use.

**Participants 15 years or younger**

If you are 15 years or younger please show this sheet to your parents. Should you decide to take part in the project your parents need to sign the Consent Form. Participants who are under 15 years of age will not be asked to provide a blood sample.

If you would like to contribute to this project, please fill out and sign the Consent Form.

**COMPENSATION**

In the unlikely event of a physical injury as a result of your participation in this study, you may be covered by ACC under the Injury Prevention, Rehabilitation and Compensation Act. ACC cover is not automatic and your case will need to be assessed by ACC according to the provisions of the 2002 Injury Prevention Rehabilitation and Compensation Act. If your claim is accepted by ACC, you still might not get any compensation. This depends on a number of factors such as whether you are an earner or non-earner. ACC usually provides only partial reimbursement of costs and expenses and there may be no lump sum compensation payable. There is no cover for
mental injury unless it is a result of physical injury. If you have ACC cover, generally this will affect your right to sue the investigators.

If you have any questions about ACC, contact your nearest ACC office or the investigator.

This study has received ethical approval from the ________ Ethics Committee.

If you have any queries or concerns regarding your rights as a participant in this study, you may wish to contact a Health and Disability Advocate, telephone 0800 555 050 Northland to Franklin.

For Māori health support, or to discuss any concerns or issues regarding this study, please contact Mata Forbes RGON, Māori Health Services Co-ordinator / Advisor, 5th Level, GM Suite, Auckland City Hospital. Tel 307 4949 extn. 23939 or Mobile 021 348 432
**Auckland Staphylococcal Study – Case**

**Demographics**
- M/F/O: Age ______
- Country of birth: NZ O Other ______
- Occupation: ________ Marital Status: M/DF O Single O Sep O
- Address: ____________________________
- Community services card O High use health card O Insurance O
- Regular GP O PHO O GP visits in past month _____ 3months_____
- Secondary school ≤3yr O 4-5 yrs O tertiary O
- Annual pre-tax income $____________________ Cell phone O
- DPB O Sickness O Invalids O War pension O Other ____________

**Case Details**
- Date of Admission ________ Discharge/Death_________
- Outcome ______________________________________________________
- Diagnosis ______________________________________________________
- Complications ______________________________________________________
- Blood cultures O Dates_________________________/________________________
- Abx ___________________________ ___/___ to ___/___
  ___________________________ ___/___ to ___/___
  ___________________________ ___/___ to ___/___
- Surgery O ________________ ___/___ _________________ ___/___
- Ix ___________________________ ___/___
- T ______ BP / PR____ RR____ Na+____ K+____
- Creat____ ARF O PCV____ WBC ___ GCS _____
- Organ Dysfxn ______
- ABG O pH_____ PO2 ___ PC02 ____ FIO2 ___ A-a ____ HCO3____
- Height ________ Weight ________ Peak CRP ______

**Exposure?**
- Household member O Prior Hospital Admit ________
- Previous staph infection: ________________ IV access O
- Prior Hospital Visit 1MO O 3 MO O
- Social group O __________________ Church group O ________________
- Sports group O ______________
- Notes: ______________________________

**Past History**
- Boils O cellulitis O Eczema O Skin Disease __________________________
- Osteomyelitis O BE O Organ abscess O RF O
- Asthma O Sinusitis O IHD O Cancer O Diabetes O INSULIN O
- Lipids O Renal O PD O CVL O Haemodialysis O __________
- Immune suppressed O ________________
- Smoker O #/day ______
- Alcohol O #drinks/week ______
- Other recreational Drug O ________________
Auckland Staphylococcal Study – Household details

MMH O    ACH O    Case # ______
Dwelling type: Detached house O    Attached house/apt O    Mobile O
Owner O    Council O    Housing NZ O    Rented O    Other _____________
Number of people _____   Ages ______________________
HCW in household O    Unemployed person O    Number of unemployed O
Any member contact with hospital in 6 months O ___________/___________
Number of bedrooms _____     How many share bedroom? ______________
Age/gender in shared rooms ____/____     ____/____     ____/____     ____/____     ____/____
Phone O        # cars ______
Home heating in winter: Electric O    Gas O    Fire O    Other O    none O
B.4 Revised consent form for carriage study

School of Medical Sciences
Department of Molecular Medicine & Pathology

CONSENT FORM

Title of project: The Auckland Staphylococcal Study

Principal Investigator: Dr. Stephen Ritchie
Dept. of Molecular Medicine & Pathology, School of Medicine
Tel.: (9) 367-0000 ext 22968

Head of Department: Prof. Peter Browett
Dept. of Molecular Medicine & Pathology, School of Medicine
Tel.: (9) 373-7599 Ext. 86218

REQUEST FOR INTERPRETER

<table>
<thead>
<tr>
<th>Language</th>
<th>Translation</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>I wish to have an interpreter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Māori</td>
<td>E hiahia ana ahau ki tetahi kaiwhakamaori/ kaiwhaka pakeha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>Ou te mana’o ia i ai se fa’amatala upu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tongan</td>
<td>Oku ou fiema’u ha fakatonulea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cook Island</td>
<td>Ka inangaro au i tetai tangata uri reo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niuean</td>
<td>Fia manako au ke fakaaoaga e taha tagata fakahokohoko kupu</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I have read and I understand the information sheet dated Nov 2007 for volunteers taking part in the Auckland Staphylococcus Study designed to investigate types of *Staphylococcus aureus* bacteria in Auckland. I have had the opportunity to discuss this study. I am satisfied with the answers I have been given.

I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time.

I understand that my participation in this study is confidential and that no material, which could identify me will be used in any reports on this study.

I have had time to consider whether to take part.

I know who to contact if I have any questions about the study.

Are you interested in the outcome of this research project? Y N

Please send a report to the following address:

I ________________________ (full name) hereby consent to take part in this study.

Signed/Date:

Signature of investigator:
Title of project: The Auckland Staphylococcal Study

Principal Investigator:
Dr Stephen Ritchie
Dept. of Molecular Medicine and Pathology, School of Medicine.
Tel.: (9) 367 0000 ext 22968

Head of Department:
Prof. Peter Browett
Dept. of Molecular Medicine and Pathology, School of Medicine.
Tel.: (9) 373-7599 Ext. 86281

Time course of study: 3 Years, until December 2009

You are being invited to take part in a research study. Your participation is entirely voluntary (your choice), and you need to decide whether or not you want to take part. Please read the following information carefully and ask the doctor any questions you may have.

Your contribution to this project

*Staphylococcus aureus* is a bacterium and is a very common cause of illness in New Zealand. It also lives in the nose of 30% of healthy members of the population (approx 400,000 people in Auckland); these people are “carriers” and very rarely get sick.

You are a healthy member of the population - we would like to take a swab from the front part of the nose of your nose to assess whether you are a “carrier” and, if so, what type of staphylococcus you have. This is part of a larger study that will enable us to compare the different types of *Staphylococcus aureus* in Auckland; for example we will compare the types found in healthy people with the types that cause infections in Auckland City and Middlemore hospitals.

All participants will be asked to complete a questionnaire which should take about 5 minutes. You do not have to answer all of the questions and you may stop the questionnaire at any time.

Benefits, risks and safety

The swab is taken from the front part of the nose and does not hurt. There will be no genetic tests performed.
There are no payments or rewards for participation in this study, but we hope to answer some important questions about an infection that is very serious and common in Auckland.

You are able to withdraw from this study at any time, without having to give a reason, and the sample you have provided will be removed from the study and destroyed.

The study will be carried out at the Auckland Medical School over the next 3 years. At the end of this period, all data will be evaluated and might be sent to a scientific journal for publication. The study will also be written up in the form of a doctoral thesis. If you are interested in the outcome of this study, please provide your address on the Consent Form and a short report will be sent to you when the data have been evaluated; this will not be possible until the analysis has been performed and may take up to 3 years.

No material, which could personally identify you will be used in any reports on this study. We do not need your name but we do require some information to help us better understand the results. The information we wish to obtain includes the following:

Your age
Your gender
Your home address
Ethnic background and country of birth

This study has received ethical approval from the Northern Y Regional Ethics Committee. The study has also has received approval from the Māori Research Review Committee at the Auckland District Health Board and the Māori Ethics Review Committee at the Counties Manukau District Health Board.

If you have any queries or concerns regarding your rights as a participant in this study, you may wish to contact a Health and Disability Advocate, telephone 0800 555 050 Northland to Franklin.

For Māori health support, or to discuss any concerns or issues regarding this study, please contact Mata Forbes RGON, Māori Health Services Co-ordinator / Advisor, 5th Level, GM Suite, Auckland City Hospital. Tel 307 4949 extn. 23939 or Mobile 021 348 432

If you would like to contribute to this project, please fill out and sign the Consent Form.
B.6 Carriage study data collection form

School of Medical Sciences
Department of Molecular Medicine & Pathology

The Auckland Staphylococcal Study - Questionnaire

How old are you? ______
Gender:  Male  O  Female  O
Country of birth:  NZ  O  Other:  ____________  How long have you lived in NZ? ____________
Address: __________________________________________________________

What ethnic group(s) do you belong to? __________________________________________

Do you have a Community services card? Yes  O  No  O
Do you have Private Health Insurance? Yes  O  No  O
Do you have a Regular GP? Yes  O  No  O
Have you seen your GP in the past 1 month? Yes  O  No  O
In the past year have you had: Boils Yes  O  No  O
Eczema Yes  O  No  O
Other Skin Disease (what?) __________________________

Have you ever had an infection with staphylococcus? Yes  O  No  O
When?

Have you been in a hospital in the past year? Yes  O  No  O
What month?
Were you admitted yourself?  O  or a Visitor?  O
Has anyone in your house stayed in hospital in the last year? Yes  O  No  O
How many people live at your home including you? ____________
How many bedrooms are there at your home? ____________
Does anyone in your house work as a healthcare worker? Yes  O  No  O
What type of work do they do? _____________________________________
References cited


125 Posada D. Using MODELTEST and PAUP* to select a model of nucleotide substitution. Curr Protoc Bioinformatics. 2003;Chapter 6:Unit 6.5.


