
Evolutionary Dynamics of Successful Clones of Methicillin-Resistant *Staphylococcus aureus* in Australia and New Zealand

Sarah Louise Baines

MSc BMLSc

ORCID: 0000-0002-0557-0518

Doctor of Philosophy

October 2019

Department of Microbiology & Immunology,
The University of Melbourne at The Peter Doherty Institute for Infection & Immunity

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

Abstract

Staphylococcus aureus is a common cause of bacterial infections in humans and a leading nosocomial pathogen, associated with significant morbidity, mortality and economic impact. A cornerstone in the evolution of staphylococcal lineages that infect humans has been their remarkable ability to rapidly and efficiently develop or acquire mechanisms of antimicrobial resistance, impacting effective disease management, prevention, and eradication. Improving these measures or developing novel approaches requires a comprehensive understanding of the infecting agent and while *S. aureus* has been extensively studied there remains considerable gaps in our knowledge surrounding this pathogen, especially concerning population dynamics. What drives the emergence and/or persistence of certain staphylococcal lineages? What evolutionary pathways and molecular mechanisms are being utilised and under what circumstances? What environmental and host factors have the greatest influence on bacterial population adaptation? And ultimately, what are the consequence of these population level changes and what impact do these have on staphylococcal disease?

This thesis represents three projects undertaken to strengthen our understanding of the evolutionary dynamics surrounding staphylococcal population adaptation, using a combination of comparative genomics and detailed phenotypic profiling to provide insight into antibiotic-resistant lineages of *S. aureus* that circulate in Australia and New Zealand.

The first project investigated the long-term persistence of the globally disseminated, multidrug resistant hospital-associated methicillin-resistant *S. aureus* (MRSA) lineage ST239 in Australia. From this work, it has been identified that the ST239 MRSA population circulating in Australia represents not one, but two genetically distinct clades; the previously unrecognised introduction of the Asian-Australian clade followed by successful local expansion in the state of Victoria has contributed to the persistence of ST239 by supplementing the diminishing population size of the local clone, the Australian clade. The ability of the Asian-Australian clade to spread following its introduction is owed to the reduced susceptibility this population developed against anti-MRSA antibiotics, namely the glycopeptides and daptomycin, prior to its importation. This same phenotype has convergently emerged in the local Australian clade. However, this adaption has come at a cost as both clades were found to have reduced replicative fitness and impaired pathogenic potential, the latter occurring through loss of functionality or reduced expression of a staphylococcal global virulence regulatory system, the accessory gene regulator.

The second project considered a different evolutionary circumstance; the rapid emergence of a novel MRSA lineage in the New Zealand community. This region has a high incidence of staphylococcal skin and soft tissue infection and over the last two decades has seen a significant shift in the local molecular epidemiology of *S. aureus*, with the emergence of multiple fusidic acid resistant clones. This work has focused on the predominant MRSA lineage, an ST5 clone locally referred to as AK3. Genomic investigation of this lineage has found that

AK3 represents a single phylogenetic clade that has arisen through local population expansion, having emerged from a resident fusidic acid susceptible methicillin-susceptible *S. aureus* upon the acquisition of a novel chimeric SCCmecIVa-*fusC* II mobile element harbouring the fusidic acid resistance determinant *fusC*. This phenomenon was not restricted to ST5, with the other newly emerged lineages having acquired *fusC* via structurally distinct SCC and SCCmec elements. Indicating that, the unregulated use of fusidic acid in the region has supported the emergence and expansion of these novel lineages and is potentially contributing to the high local incidence of staphylococcal infection.

Mobile elements can greatly influence staphylococcal populations, therefore the third project focused on exploring the evolution of the multidrug resistant plasmid family pSK1 and the role it has played in augmenting antimicrobial resistance and biocide tolerance in the ST239 MRSA population in Australia. Modelling the evolutionary history of this plasmid identified that it emerged during the late 1970s and over the following four decades has undergone significant structural change, involving a combination of chromosomal integration, transposon loss/gain, structural inversion and deletion events. When aligned to a phylogenetic model of the ST239 population, it became apparent that these changes in plasmid configuration represented a clear pathway of step-wise adaptation; the shortened, chromosomally integrated plasmid structural variants having emerged on multiple occasions in a convergent manner. Further, these changes correlated with the development of enhanced tolerance against chlorhexidine, an important finding as it implicates biocide use as a factor potentially influencing MRSA evolution.

Collectively, this work enhances our understanding of how antibiotic-resistant lineages of *S. aureus* evolve and adapt; exemplifying the evolutionary pathways facilitating adaptation in staphylococcal populations and the circumstances under which they are used. Further, it provides detailed insight into two highly successful lineages of MRSA. This knowledge can be exploited to improve measures that reduce the burden of antibiotic-resistant staphylococcal infection. Importantly, this work enhances and reinforces our awareness about the consequences of antimicrobial overuse and misuse.

Declaration

This is to certify that:

- I. This thesis, entitled “Evolutionary Dynamics of Successful Clones of Methicillin-Resistant *Staphylococcus aureus* in Australia and New Zealand”, comprises only my original work towards this Ph.D except where otherwise indicated;
- II. Due acknowledgement has been made in the text to all other material used;
- III. The thesis is fewer than 100,000 words in length, exclusive of tables, figures, figure legends, bibliography, and appendices.

.....

Sarah Louise Baines MSc BMLSc

Department of Microbiology & Immunology

The University of Melbourne at The Peter Doherty Institute for Infection & Immunity

Preface

This preface provides a summary of the contents of each chapter in this thesis. The chapters that include publications are indicated. The nature and extent of the contributions to each chapter made by the author of this thesis are listed.

Funding: Sarah Louise Baines was supported by an Australian Government Research Training Program Scholarship. Additional funding was awarded in the form of a Victorian Fellowship, provided by the Victoria State Government to undertake an overseas study mission. This supported the research presented in this thesis and unpublished projects representing extensions of this work. Funding awarded to co-authors of the publications presented in this thesis are indicated below.

Chapter 1: Introduction is an original overview of the background, key concepts and motivations for this thesis, with the original works cited.

Chapter 2: Convergent Adaptation in the Dominant Global Hospital Clone ST239 of Methicillin-Resistant *Staphylococcus aureus* is an original work that resulted in the following publication:

Baines, S. L., K. E. Holt, M. B. Schultz, T. Seemann, B. O. Howden, S. O. Jensen, S. J. van Hal, G. W. Coombs, N. Firth, D. R. Powell, T. P. Stinear and B. P. Howden (2015). "Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*." MBio; DOI: 10.1128/mBio.00080-15.

The author of this thesis was the first author and main contributor of the work presented in this publication. The nature and extent of my contributions to this chapter are described below.

- I contributed to the generation of data (exceptions listed below), interpretation of results, and development of the concepts described in the publication with BPH and TPS.
- I generated all new sequence data presented in this publication. Several bacterial isolates and/or whole genome sequence data was kindly provided by SOJ, SJH and NF.
- I performed all bioinformatics analyses, except for recombination detection which was performed by KEH. Supervision, teaching and/or software was provided by TPS, TS, KEH, MBS and DRP.
- I performed all phenotypic experiments. Assistance with a murine sepsis model was provided by BOH. All animal studies were performed in accordance with the Animal Research Ethics Committee at The University of Melbourne (Ethics Application #1212591).

- I was responsible for the planning, drafting, and editing of the publication with BPH and TPS.
- All co-authors reviewed and edited the manuscript.

Funding associated with this publication include National Health and Medical Research Council, Australia: Fellowships awarded to BPH (APP1023526), TPS (APP1008549), and KEH (APP1061409).

Chapter 3: Rapid Emergence and Evolution of *Staphylococcus aureus* Clones Harboursing *fusC*-Containing Staphylococcal Cassette Chromosome Elements is an original work that resulted in the following publication:

Baines, S. L., B. P. Howden, H. Heffernan, T. P. Stinear, G. P. Carter, T. Seemann, J. C. Kwong, S. R. Ritchie and D. A. Williamson (2016). "Rapid emergence and evolution of *Staphylococcus aureus* clones harboring *fusC*-containing staphylococcal cassette chromosome elements." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.03020-15.

The author of this thesis was the first author and main contributor of the work presented in this publication. The nature and extent of my contributions to this chapter are described below.

- I contributed to the generation of data (exceptions listed below), interpretation of results, and development of the concepts described in the publication with DAW, BPH and TPS.
- Isolate collection and initial phenotypic characterisation were performed by staff at the Institute of Environmental Science and Research, New Zealand. ESR staff also performed extraction of genomic DNA, with sequencing conducted at the J. Craig Venter Institute, CA, United States.
- I performed the genomic DNA extraction for long-read sequencing, with sequencing conducted at the Duke Center for Genomic and Computational Biology, NC, United States.
- I performed all bioinformatics analyses. Supervision, teaching and/or software was provided by TPS, TS and JCK.
- I was responsible for the planning, drafting, and editing of the publication with DAW, BPH and TPS.
- All co-authors reviewed and edited the manuscript.

Chapter 4: Remodelling of pSK1 Family Plasmids and Enhanced Chlorhexidine Tolerance in a Dominant Hospital Lineage of Methicillin-Resistant *Staphylococcus aureus* is an original work that resulted in the following publication:

Baines, S. L., S. O. Jensen, N. Firth, A. Gonçalves da Silva, T. Seemann, G. P. Carter, D. A. Williamson, B. P. Howden and T. P. Stinear (2019). "Remodelling of pSK1 family plasmids and enhanced chlorhexidine tolerance in a dominant hospital lineage of methicillin-resistant *Staphylococcus aureus*." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.02356-18.

The author of this thesis was the first author and main contributor of the work presented in this publication. The nature and extent of my contributions to this chapter are described below.

- I contributed to the design of the study, generation of data (exceptions listed below), interpretation of results, and development of the concepts described in the chapter with BPH and TPS.
- Bacterial isolates collected as part of the Australian and New Zealand Cooperative in Staphylococcal Sepsis and the Vancomycin Efficacy in Staphylococcal Sepsis studies were kindly provided by Natasha E. Holmes.
- I performed the genomic DNA extraction for long-read sequencing. Sequencing was performed with assistance from GPC.
- I performed all bioinformatics analyses. Supervision, teaching and/or software was provided by TPS, TS and AGS.
- I performed all phenotypic experiments with assistance from Eloise Alison.
- Assistance was provided by NF and SOJ in the interpretation of likely evolutionary mechanisms underlying changes in plasmid configuration.
- I was responsible for the planning, drafting, and editing of the manuscript with BPH and TPS.
- All co-authors reviewed and edited the manuscript.

Funding associated with this publication include National Health and Medical Research Council, Australia: Fellowship awarded to BPH (GNT1105905), Centres of Research Excellence grant awarded to BPH (GNT1102962), and Project Grant awarded to NF and SOJ (GNT457454).

Chapter 5: Conclusions is an original summary of the significance, implications and limitations of the work presented in this thesis.

Additional publications arising during candidature

Berti, A. D., **S. L. Baines**, B. P. Howden, G. Sakoulas, V. Nizet, R. A. Proctor and W. E. Rose (2015). "Heterogeneity of genetic pathways toward daptomycin nonsusceptibility in *Staphylococcus aureus* determined by adjunctive antibiotics." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.04990-14.

Carter, G. P., A. H. Buultjens, S. A. Ballard, **S. L. Baines**, T. Tomita, J. Strachan, P. D. Johnson, J. K. Ferguson, T. Seemann, T. P. Stinear and B. P. Howden (2016). "Emergence of endemic MLST non-typeable vancomycin-resistant *Enterococcus faecium*." *Journal of Antimicrobial Chemotherapy*; DOI: 10.1093/jac/dkw314

Baines, S. L.,* D. A. Williamson,* G. P. Carter, A. Gonçalves da Silva, X. Ren, J. Sherwood, M. Dufour, M. B. Schultz, N. P. French, T. Seemann, T. P. Stinear and B. P. Howden (2016). "Genomic insights into a sustained national outbreak of *Yersinia pseudotuberculosis*." *Genome Biology and Evolution*; DOI: 10.1093/gbe/evw285.

**These authors contributed equally to this work.*

Kpeli, G., I. Darko Otchere, A. Lamelas, A. L. Buultjens, D. Bulach, **S. L. Baines**, T. Seemann, S. G. Giulieri, Z. Nakobu, S. Y. Aboagye, E. Owusu-Mireku, G. Pluschke, T. P. Stinear and D. Yeboah-Manu (2016). "Possible healthcare-associated transmission as a cause of secondary infection and population structure of *Staphylococcus aureus* isolates from two wound treatment centres in Ghana." *New Microbes and New Infections*; DOI: 10.1016/j.nmni.2016.07.001.

Kpeli, G., A. H. Buultjens, S. G. Giulieri, E. Owusu-Mireku, S. Y. Aboagye, **S. L. Baines**, T. Seemann, D. Bulach, A. Gonçalves da Silva, I. R. Monk, B. P. Howden, G. Pluschke, D. Yeboah-Manu and T. P. Stinear (2017). "Genomic analysis of ST88 community-acquired methicillin resistant *Staphylococcus aureus* in Ghana." *PeerJ*; DOI: 10.7717/peerj.3047.

Gonçalves da Silva, A., **S. L. Baines**, G. P. Carter, H. Heffernan, N. P. French, X. Ren, T. Seemann, D. Bulach, J. C. Kwong, T. P. Stinear, B. P. Howden and D. A. Williamson (2017). "A phylogenomic framework for assessing the global emergence and evolution of clonal complex 398 methicillin-resistant *Staphylococcus aureus*." *Microbial Genomics*; DOI: 10.1099/mgen.0.000105.

Buultjens, A. H., K. Y. L. Chua, **S. L. Baines**, J. Kwong, W. Gao, Z. Cutcher, S. Adcock, S. A. Ballard, M. B. Schultz, T. Tomita, N. Subasinghe, G. P. Carter, S. J. Pidot, L. Franklin, T. Seemann, A. Gonçalves Da Silva, B. P. Howden and T. P. Stinear (2017). "A supervised statistical learning approach for accurate *Legionella pneumophila* source attribution during outbreaks." *Applied and Environmental Microbiology*; DOI: 10.1128/AEM.01482-17.

Cameron, D. R., Y. H. Lin, S. Trouillet-Assant, V. Tafani, X. Kostoulas, E. Mouhtouris, N. Skinner, K. Visvanathan, **S. L. Baines**, B. P. Howden, I. R. Monk, F. Laurent, T. P. Stinear, B. P. Howden and A. Y. Peleg (2017). "Vancomycin-intermediate *Staphylococcus aureus* isolates are attenuated for virulence when compared with susceptible progenitors." *Clinical Microbiology and Infection*; DOI: 10.1016/j.cmi.2017.03.027

Amissah, N. A., A. H. Buultjens, A. Ablordey, L. van Dam, A. Opoku-Ware, **S. L. Baines**, D. Bulach, C. S. Tetteh, I. Prah, T. S. van der Werf, A. W. Friedrich, T. Seemann, J. M. van Dijk, Y. Stienstra, T. P. Stinear and J. W. Rossen (2017). "Methicillin resistant *Staphylococcus aureus* transmission in a Ghanaian burn unit: The importance of active surveillance in resource-limited settings." *Frontiers in Microbiology*; DOI: 10.3389/fmicb.2017.01906.

Carter, G. P., M. B. Schultz, **S. L. Baines**, A. Gonçalves da Silva, H. Heffernan, A. Tiong, P. H. Pham, I. R. Monk, T. P. Stinear, B. P. Howden and D. A. Williamson (2018). "Topical antibiotic use co-selects for the carriage of mobile genetic elements conferring resistance to unrelated antimicrobials in *Staphylococcus aureus*." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.02000-17.

Kwong, J. C., C. Lane, F. Romanes, A. Gonçalves da Silva, M. Easton, K. Cronin, M. J. Waters, T. Tomita, K. Stevens, M. B. Schultz, **S. L. Baines**, N. L. Sherry, G. P. Carter, A. Mu, M. Sait, S. A. Ballard, T. Seemann, T. P. Stinear and B. P. Howden (2018). "Translating genomics into practice for real-time surveillance and response to carbapenemase-producing *Enterobacteriaceae*: evidence from a complex multi-institutional KPC outbreak." *PeerJ*; DOI: 10.7717/peerj.4210.

Sherry, N. L., **S. L. Baines** and B. P. Howden TP (2018). "Ceftazidime/avibactam susceptibility by three different susceptibility testing methods in carbapenemase-producing Gram-negative bacteria from Australia." *International Journal of Antimicrobial Agents*; DOI: 10.1016/j.ijantimicag.2018.02.017.

Berti, A. D., N. Shukla, A. D. Rottier, J. S. McCrone, H. M. Turner, I. R. Monk, **S. L. Baines**, B. P. Howden, R. A. Proctor and W. E. Rose (2018). "Daptomycin selects for genetic and phenotypic adaptations leading to antibiotic tolerance in MRSA." *Journal of Antimicrobial Chemotherapy*; DOI: 10.1093/jac/dky148.

Giulieri, S. G., **S. L. Baines**, R. Guerillot, T. Seemann, A. Gonçalves da Silva, M. B. Schultz, R. C. Massey, N. E. Holmes, T. P. Stinear and B. P. Howden (2018). "Genomic exploration of sequential clinical isolates reveals a distinctive molecular signature of persistent *Staphylococcus aureus* bacteraemia." *Genome Medicine*; DOI: 10.1186/s13073-018-0574-x.

Carter, G. P., J. E. Ussher, A. Gonçalves da Silva, **S. L. Baines**, H. Heffernan, T. V. Riley, R. Broadbent, A. van der Linden, J. Lee, I. R. Monk, T. P. Stinear, B. P. Howden and D. A. Williamson (2018). "Genomic analysis of multi-resistant *Staphylococcus capitis* associated with neonatal." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.00898-18.

Guerillot, R., L. Li, **S. L. Baines**, B. Howden, M. B. Schultz, T. Seemann, I. R. Monk, S. J. Pidot, W. Gao, S. G. Giulieri, A. Gonçalves da Silva, A. D'Agata, T. Tomita, A. Y. Peleg, T. P. Stinear and B. P. Howden (2018). "Comprehensive antibiotic-linked mutation assessment by resistance mutation sequencing (RM-seq)." *Genome Medicine*; DOI: 10.1186/s13073-018-0572-z.

Lee, R. S., A. Gonçalves da Silva, **S. L. Baines**, J. Strachan, S. Ballard, G. P. Carter, J. C. Kwong, M. B. Schultz, D. M. Bulach, T. Seemann, T. P. Stinear and B. P. Howden (2018). "The changing landscape of vancomycin-resistant *Enterococcus faecium* in Australia: a population-level genomic study." *Journal of Antimicrobial Chemotherapy*; DOI: 10.1093/jac/dky331.

Bouchez, V., **S. L. Baines**, S. Guillot and S. Brisse (2018). "Complete genome sequences of *Bordetella pertussis* clinical isolate FR5810 and reference strain Tohama from combined Oxford Nanopore and Illumina sequencing." *Microbial Resource Announcements*; DOI: 10.1128/MRA.01207-18.

Ingle, D. J., M. Easton, M. Valcanis, T. Seemann, J. C. Kwong, N. Stephens, G. P. Carter, A. Gonçalves da Silva, J. Adamopoulos, **S. L. Baines**, K. E. Holt, E. P. F. Chow, C. K. Fairley, M. Y. Chen, M. Kirk, B. P. Howden and D. A. Williamson (2019). "Co-circulation of multidrug-resistant *Shigella* among men who have sex with men, Australia." *Clinical Infectious Diseases*; DOI: 10.1093/cid/ciz005.

Jiang, J. H., C. Dexter, D. R. Cameron, I. R. Monk, **S. L. Baines**, I. J. Abbott, D. W. Spelman, X. Kostoulis, C. Nethercott, B. P. Howden and A. Y. Peleg (2019). "Evolution of daptomycin resistance in coagulase negative staphylococci involves mutations of the essential two component regulator WalKR." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.01926-18.

Minhas, V., R. M. Harvey, L. J. McAllister, T. Seemann, A. E. Syme, **S. L. Baines**, J. C. Paton and C. Trappetti (2019). "Capacity to utilize raffinose dictates pneumococcal disease phenotype." *mBio*; DOI: 10.1099/jmm.0.000916.

Douglas, A. P., C. Marshall, **S. L. Baines**, D. Ritchie, J. Szer, V. Madigan, H. T. Chan, S. A. Ballard, B. P. Howden, K. Buising and M. A. Slavin (2019). "Utilising genomic analyses to investigate the first outbreak of *vanA* vancomycin-resistant *Enterococcus* in Australia with emergence of daptomycin non-susceptibility." *Journal of Medical Microbiology*; DOI: 10.1099/jmm.0.000916.

Bojang, A., **S. L. Baines**, L. Donovan, R. Guerillot, K. Stevens, C. Higgs, C. Bottomley, O. Secka, M. B. Schultz, A. Gonçalves da Silva, T. Seemann, T. P. Stinear, A. Roca and B. P. Howden (2019). "Genomic investigation of *Staphylococcus aureus* recovered from Gambian women and newborns following an oral dose of intra-partum azithromycin." *Journal of Antimicrobial Chemotherapy*. DOI: 10.1093/jac/dkz341.

Coombs, G. W., **S. L. Baines**, B. P. Howden, K. M. Swenson and F. G. O'Brien (2020). "Diversity of bacteriophages encoding Panton-Valentine leukocidin in temporally and geographically related *Staphylococcus aureus*." *PLoS One*. DOI: 10.1371/journal.pone.0228676.

Mangas, K. M., A. H. Buultjens, J. L. Porter, **S. L. Baines**, E. Marion, L. Marsollier, N. J. Tobias, S. J. Pidot, K. M. Quinn, D. J. Price, K. Kedzierska, W. Zeng, D. C. Jackson, B. Y. Chua* and T. P. Stinear* (2020). "Vaccine-specific immune responses against *Mycobacterium ulcerans* infection in a low-dose murine challenge model." *Infection and Immunity*. DOI: 10.1128/IAI.00753-19. *These authors contributed equally to this work.

Available as a preprint

Baines, S. L., A. Gonçalves da Silva, G. Carter, A. V. Jennison, I. Rathnayake, R. M. Graham, V. Sintchenko, Q. Wang, R. J. Rockett, V. J. Timms, E. Martinez, S. Ballard, T. Tomita, N. Isles, K. A. Horan, W. Pitchers, T. P. Stinear, D. A. Williamson, B. P. Howden, Communicable Diseases Genomics Network (CDGN), T. Seemann (2019). "Complete microbial genomes for public health in Australia and Southwest Pacific." *bioRxiv*. DOI: 10.1101/829663.

Under review or in preparation

Jenson, R. E., **S. L. Baines**, B. P. Howden, N. N. Mishra, S. Farah, C. Lew, A. D. Berti, S. K. Shukla, A. S. Bayer and W. E. Rose (2019). "Prolonged exposure to beta-lactam antibiotics reestablishes susceptibility of daptomycin-nonsusceptible *Staphylococcus aureus* to daptomycin."

Bojang, A. **S. L. Baines**, B. Camara, R. Guerillot, R. S. Marquez, O. Secka, U. D'Alessandro, C. Bottomley, B. P. Howden and A. Roca (2019). "Impact of oral intrapartum azithromycin on the nasopharyngeal staphylococci macrolide resistance of babies: a randomised control trial."

Baines, S. L.,* C. Miller,* S. Dey,* P. Smolenski,* P. Kulkarni* and A. Berti (2019). "Suppression of daptomycin resistance development in *Staphylococcus aureus* is a class effect of beta lactams and is independent of daptomycin-beta lactam synergy." *These authors contributed equally to this work.

Acknowledgements

I am sincerely grateful for the assistance and support of numerous individuals, for the work presented in this thesis (and the maintenance of my mental sanity) would not have been possible without them.

First and foremost, I would like to thank my supervisors Ben Howden and Tim Stinear. Their continual support throughout my MSc and then Ph.D has meant a lot. Both have strongly encouraged me to follow the research that has genuinely fascinated me, even when there was no clear endpoint (or at least not one that I saw at the time I proposed the idea), and they have always been generous in providing their time and advice when needed. Also, the many many many side projects I was allowed to undertake simply because I was interested.

I would also like to thank the other members of my Ph.D committee, Kat Holt and Roy Robins-Browne, for their invaluable assistance, regularly helping me to see my research from a different perspective. I would especially like to thank Roy for his “life” lessons and very practical advice on what it takes to thrive in science.

I owe a big thank you to the many members of the Howden, Stinear and Williamson labs, all of who have helped me at some point during this time. Jean Lee, Ian Monk, Wei Gao, Stefano Giulieri, Romain Guerillot, Natasha Holmes, Kyra Chua and Ya-Hsun Lin – for your collective encyclopedia-level knowledge on *Staphylococcus* and your willingness to share that knowledge with me; Sacha Pidot – for fielding many of my non-staphylococcal questions and always having an answer or idea; Jess Porter – for making sure I always had the reagents I needed (apologies for my lack of organisation some days) and putting up with my many questions about where to find equipment after every reorganisation of the lab; Torsten Seemann, Anders Gonçalves da Silva and Jason Kwong – thank you for all the support (not just bioinformatic). You are all people I felt I could approach with any idea, no matter how computationally unrealistic, and I owe so much of my bioinformatics knowledge to you; Kirstie Mangas and Andrew Buultjens – I consider myself lucky to have had both you undertaking your Honours and then Ph.D at the same time (we are all nearly there!). And to Kirstie – thank you for the many times that a single comment turned into an hour-long chat; Danielle Ingle, Claire Gorrie, Jake Lacey and Andre Mu – thank you for being the people I could go to with any question no matter how basic or out of the box, for being my sounding-board, for all the coffee runs and evenings at the pub, and for being the “sensible” voices when I lost perspective. I only hope I was able to be the same for you.

Thank you to the staff at MDU PHL (and the Austin during my MSc) who kindly put up with me when I was searching for “historic” samples or wanting to use equipment. In particular I would like to mention Susan Ballard, who I have been lucky to work with since I first started at the Austin. You are not only someone I go to for scientific advice but over the years you’ve become a good friend and mentor. I can’t express how much I appreciate all our chats and your advice.

From the Dept. of Microbiology & Immunology, thank you to the teaching team (Helen Cain, Karena Waller, Louise Adams, Odilia Wijburg and Cheryl Power) for offering me so many opportunities to be involved with the

undergraduate and graduate microbiology courses. I have learnt so much from those experiences, one of which is a love for teaching. A big thank you to Rebecca Whitsed for all the admin support and going above and beyond, regularly checking in to make sure everything was ok.

Thank you to the Australian Society for Microbiology for the chance to serve on the Victorian Branch committee. This has been a wonderful opportunity for me to give back to the microbiology community, and one I hope to continue.

Lastly, I would like to thank my family. Mum and Dad - thank you for your unwavering support and willingness to listen to me talk about my research for hours on end. You recognised my love for science and research long before I did; Emily – thanks for being the sister that was always happy to provide a “reality check” and for making me laugh when I needed it most, even when it was at my expense.

Table of Contents

Abstract	II
Declaration	IV
Preface	V
Additional publications arising during candidature.....	VIII
Acknowledgements	XII
Table of Contents.....	XIV
List of Tables & Figures.....	XVII
Abbreviations	XIX
 <u>Chapter 1</u>.....	 1
1.1 The Staphylococci	2
1.2 The Pathogen <i>S. aureus</i>	2
1.3 The Genome of <i>S. aureus</i>.....	4
1.3.1 Lineage Classification	4
1.4 Antimicrobial Resistance in <i>S. aureus</i>.....	6
1.4.1 Evolutionary History of Antibiotic Resistant <i>S. aureus</i>	7
<i>First Wave: Penicillin-Resistant S. aureus</i>	<i>7</i>
<i>Second Wave: Methicillin-Resistant S. aureus.....</i>	<i>9</i>
<i>Third Wave: Healthcare-Associated MRSA.....</i>	<i>9</i>
<i>Fourth Wave: Community-Associated MRSA</i>	<i>10</i>
1.4.2 Ongoing Challenge of Glycopeptide and Daptomycin Resistance	11
<i>Glycopeptides & Lipoglycopeptides</i>	<i>11</i>
<i>Daptomycin.....</i>	<i>13</i>
1.4.3 Emerging Challenge of Topical Antibiotic and Disinfectant Resistance	14
<i>Fusidic Acid</i>	<i>14</i>
<i>Chlorhexidine.....</i>	<i>16</i>
1.5 Mechanisms of Virulence in <i>S. aureus</i>	20
1.5.1 Regulating Virulence: The Accessory Gene Regulator	22
1.6 Mobile Genetic Elements in <i>S. aureus</i>.....	24
1.6.1 Staphylococcal Cassette Chromosome	24
1.6.2 pSK1 Plasmid Family.....	26
1.7 Research Objectives.....	28

Chapter 2	29
2.1 Introduction	30
2.1.1 Project Aims	31
2.1.2 Chapter Overview	31
2.2 Results	31
2.2.1 Publication: Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant <i>Staphylococcus aureus</i>	31
2.2.2 Supplementary Materials	41
Text S1: Supplementary Materials & Methods	41
2.3 Extended Analysis	50
2.3.1 Materials & Methods	50
2.3.2 Results & Discussion	53
2.4 Conclusion	59
 Chapter 3	 60
3.1 Introduction	61
3.1.1 Project Aims	62
3.1.2 Chapter Overview	62
3.2 Results	62
3.2.1 Publication: Rapid emergence and evolution of <i>Staphylococcus aureus</i> clones harbouring <i>fusC</i> -containing staphylococcal cassette chromosome elements	62
3.2.2 Supplementary Material	70
3.3 Extended Analysis	71
3.3.1 Materials & Methods	71
3.3.2 Results & Discussion	73
3.4 Conclusion	77
 Chapter 4	 78
4.1 Introduction	79
4.1.1 Project Aims	80
4.1.2 Chapter Overview	80
4.2 Results	80
4.2.1 Publication: Remodelling of pSK1 family plasmids and enhanced chlorhexidine tolerance in a dominant hospital lineage of methicillin-resistant <i>Staphylococcus aureus</i>	80
4.2.2 Supplementary Material	97
Supplementary Results: Part A. Description of the pSK1-like plasmid variants	97
Supplementary Results: Part B. Investigation of the association between plasmid evolution and biocide tolerance	99
Supplementary Materials & Methods	101
4.3 Summary	111

<u>Chapter 5</u>	112
5.1 Introduction	113
5.2 Key Findings & Significance	113
5.2.1 Emergence & Establishment of Novel Lineages	113
<i>Lineage emergence through gene acquisition</i>	114
<i>Lineage emergence through introduction into a new environment</i>	114
5.2.2 Adaptation to Promote Lineage Persistence	115
<i>Gene acquisition to maintain and/or increase population genetic diversity</i>	115
<i>Increasing genetic diversity through co-selection</i>	116
<i>Positive selection of resistance-associated genotypes</i>	116
<i>Consequences of adaptation: trading virulence for resistance</i>	117
5.2.3 Lineage Disappearance	117
 <u>Appendix 1</u>	119
<u>Appendix 2</u>	131
<u>Appendix 3</u>	140
 <u>Bibliography</u>	152

List of Tables & Figures

Chapter 1

Figure 1.1 Genetic diversity of <i>S. aureus</i>	5
Figure 1.2 Timeline of the development and/or clinical use of antimicrobial agents and the <i>in vivo</i> emergence of resistance in <i>S. aureus</i>	6
Figure 1.3 Illustration of late stage peptidoglycan biosynthesis in <i>S. aureus</i> and the mechanisms targeted by cell wall active antimicrobial agents	8
Figure 1.4 Illustration of ribosomal peptide synthesis, mode of action of fusidic acid and mechanisms of resistance	15
Table 1.1 Summary of antimicrobial resistance mechanisms in <i>S. aureus</i>	17
Figure 1.5 The accessory gene regulator	23
Figure 1.6 Structure of staphylococcal cassette chromosome elements	25

Chapter 2

Publication figures & tables

FIG 1 Global phylogeny and molecular clock of ST239 MRSA.....	34
FIG 2 Genomic characterization of Australian ST239 clades	35
TABLE 1 Comparison of the antimicrobial susceptibility and bacterial fitness of temporal and phylogenetic groups	36
FIG 3 GWAS of reduced vancomycin susceptibility in ST239 MRSA.....	36
FIG 4 Evolution of ST239 MRSA toward attenuated virulence.....	38
FIG S1 Comparison of the global phylogeny of ST239 MRSA inferred by Bayesian and ML methods.....	45
FIG S2 Phenotypic characterization of Australian ST239 <i>S. aureus</i> strains.....	46
FIG S3 Increasing antimicrobial resistance of ST239 MRSA	47
FIG S4 Evolution of ST239 <i>S. aureus</i> toward attenuated virulence	48
TABLE S1 WalKR mutations in Australian ST239 MRSA.....	49
TABLE S2 Single nucleotide polymorphisms detected within the <i>agr</i> locus.....	49

Extended analysis

Figure 2.1 Global phylogeny of ST239 MRSA	54
Figure 2.2 Population structure of the Australasian lineage	55
Figure 2.3 Population structure of the Asian-Australasian lineage.....	56
Figure 2.4 Topological variation in the global phylogeny of ST239 MRSA.....	58
Figure 2.5 Fluctuations in effective population size overtime.....	58

Chapter 3

Publication figures

FIG 1 Sequence comparison of all complete genomes of clonal complex 5 (CC5) <i>S. aureus</i> strains	65
FIG 2 Population structure of fusidic acid-resistant CC5 <i>S. aureus</i>	66
FIG 3 Comparison of <i>fusC</i> -containing staphylococcal cassette chromosome elements in ST1 and ST5 lineages of <i>S. aureus</i>	67
FIG S1 Alignment of region surrounding <i>orfX</i>	70

Extended analysis

Figure 3.1 Global phylogeny of CC5 <i>S. aureus</i>	74
Figure 3.2 Comparison of the population structure of fusidic acid resistant CC5 MRSA and SCC ₄₇₆	75
Table 3.1 Classification of <i>fusC</i> -harbouring CC5 <i>S. aureus</i> in this collection	77

Chapter 4

Publication figures & tables

FIG 1 Global population structure of ST239 <i>S. aureus</i> isolates and prevalence of pSK1-like plasmids	84
FIG 2 pSK1-like structural variants	85
FIG 3 Evolutionary history of pSK1-like plasmids in the Australian ST239 clade	87
TABLE 1 Population distributions of antimicrobial and biocide resistance genes and phenotypic resistance profiles	89
TABLE 2 Investigation of chlorhexidine tolerance in Australian ST239 MRSA	90
FIG 4 Phenotypic variation in chlorhexidine tolerance	91
FIG 5 Bayesian phylogenetic model associating chlorhexidine tolerance with pSK1-like plasmid evolution	92
FIG S1 pSK1 plasmid gene presence and synteny	105
FIG S2 Structure of pSK1, pTW20_1 and pBPH2003 plasmids	106
FIG S3 Modelling temporal-association in phenotypic susceptibility data	107
FIG S4 Exploration of gene presence/absence signatures associated with phenotypic chlorhexidine tolerance	108
FIG S5 Exploration of mutations associated with phenotypic chlorhexidine tolerance	109
TABLE S1 Investigation of gentamicin resistance in Australian ST239 MRSA	110
TABLE S2 Gentamicin resistance and chlorhexidine tolerance profiles	110

Abbreviations

AAC	N-Acetyltransferase (aminoglycoside modifying enzyme)
AGR	Accessory gene regulator (system)
AME	Aminoglycoside modifying enzyme
ANT	O-Adenyltransferase (aminoglycoside modifying enzyme)
APH	O-Phosphotransferase (aminoglycoside modifying enzyme)
ASSOP	Australian Staphylococcal Sepsis Outcome Program
CA	Community-associated
CC	Clonal complex
CDS	(Predicted) protein coding sequence(s)
CHX	Chlorhexidine
D-ala-D-ala	D-alanyl[4]-D-alanine[5] (pentapeptide bridge precursor)
D-ala-D-lac	D-alanyl[4]-D-lactate[5] (pentapeptide bridge precursor)
DAPC	Discriminant analysis of principal components
DFR	Dihydrofolate reductase
DR	Direct repeat(s)
EF-G	Elongation factor G
EPS	Effective population size
FA	Fusidic acid
GTR	General time reversible (nucleotide substitution model)
GWAS	Genome wide association study
HA	Healthcare-associated
HDP	(Cationic antimicrobial) host defence peptide(s)
HGT	Horizontal gene transfer
HPD	Highest posterior density
IR	Inverted repeat(s)
IS	Insertion sequence(s)
Kb	Kilo-bases
Mb	Mega-bases
MBC	Minimum bactericidal concentration
MCC	Maximum clade credibility (tree)
MCMC	Monte Carlo Markov Chain
MCRA	Most recent common ancestor
MDS	Multidimensional scaling (plot)
MFS	Major facilitator (protein) superfamily
MGE	Mobile genetic element(s)

MIC	Minimum inhibitory concentration
ML	Maximum likelihood
MLEE	Multi-locus enzyme electrophoresis
MLST	Multi-locus sequence typing
MNP	Multi-nucleotide polymorphism(s)
MSCRAMM	Microbial surface components recognising adhesive matrix molecule
NCR	(Insertion sequence) non-coding region(s)
PFGE	Pulse-field gel electrophoresis
PBP	Penicillin binding protein
PVL	Panton-Valentine leukocidin
QAC	Quaternary ammonium compound
QRDR	Quinolone resistance determining region
RFLP	Restriction fragment length polymorphism
RRDR	Rifampicin resistance determining region
SAB	<i>Staphylococcus aureus</i> bacteraemia
SCC	Staphylococcal cassette chromosome
SMR	Small multidrug resistance (protein family)
SNP	Single-nucleotide polymorphism(s)
SSTI	Skin and soft tissue infection
ST	Sequence type
SV	(Plasmid) structural variant(s)
TIR	Terminal inverted repeat(s)
Tn	Transposon
tnp	Transposase (enzyme)
TSD	(Insertion sequence) target site duplication(s)
UCLN	Uncorrelated relaxed log-normal (molecular clock model)
VF	Virulence factor
WGS	Whole genome sequence (sequencing)

Nomenclature for antibiotic resistant S. aureus

PRSA	Penicillin-resistant <i>Staphylococcus aureus</i>
PSSA	Penicillin-susceptible <i>Staphylococcus aureus</i>
BORSA	Borderline oxacillin-resistant <i>Staphylococcus aureus</i>
MODSA	Modified (oxacillin-resistant) <i>Staphylococcus aureus</i>
MRSA	Methicillin-resistant <i>Staphylococcus aureus</i>
MSSA	Methicillin-susceptible <i>Staphylococcus aureus</i>
hVISA	Heterogenous vancomycin-intermediate <i>Staphylococcus aureus</i>
VISA	Vancomycin-intermediate <i>Staphylococcus aureus</i>
VRSA	Vancomycin-resistant <i>Staphylococcus aureus</i>
DNS	Daptomycin non-susceptible <i>Staphylococcus aureus</i>

Chapter 1

Introduction & Literature Review

Staphylococcus aureus

1.1 The Staphylococci

The genus *Staphylococcus* represents more than 70 species and subspecies of Gram-positive, predominantly catalase positive, non-motile, facultative or strict anaerobic bacteria; distinguished by their spherical shape and ability to divide in multiple planes, resulting in the formation of clusters of cocci (1, 2). Historically, the staphylococci were classified alongside the micrococci in the family *Micrococcaceae* because of their shared morphology and ability to breakdown hydrogen peroxide. The introduction of molecular biology has seen the staphylococci reclassified, first being added to the *Bacillus-Lactobacillus-Streptococcus* cluster and in 2004 designated to the family *Staphylococcaceae*, alongside other genera which share similar molecular, phenotypic, and biochemical characteristics (3-5).

Many staphylococcal species are common colonisers of the skin and mucus membranes in mammals and birds, only causing disease when innate immune barriers are compromised (6), such as disruption to the cutaneous epithelium, be through trauma or medical intervention. The most common staphylococcal species to colonise humans is *S. epidermidis*, however, the majority of staphylococcal disease in humans is caused by *S. aureus*. Commonly known as “Golden Staph”, *S. aureus* is primarily a commensal of the skin, nares and perineum, and carried by approximately one third of the human population at any given point in time, with carriage being predominantly transient (7-10). In most cases, the interaction between *S. aureus* and humans is benign (11). Nonetheless, *S. aureus* maintains the capacity to cause disease when innate immune barriers are breached and has been a recognised leading cause of bacterial infections in humans since it was first described in the late 19th century (12), with significant attributed morbidity and mortality (13, 14).

As a result of its importance in human disease, *S. aureus* has been extensively studied and the body of research concerning this pathogen is immense. The information provided in Chapter 1 is not intended to provide an exhaustive account of the literature surrounding *S. aureus* but rather a synopsis of the background knowledge relevant to this bacterium. Subsequently, provided is an overview of key concepts in staphylococcal research, namely pathogenicity and virulence, antimicrobial resistance, the staphylococcal genome, and mobile genetic elements. Following these are more detailed discussions of specific topics that are relevance to the work presented in Chapters 2, 3 and 4.

1.2 The Pathogen *S. aureus*

Staphylococcal disease can present as a wide array of clinical conditions with varying disease severity (10, 15, 16). These range from mild and self-limiting skin and soft tissue infections [SSTI] (eg. boils, folliculitis, cellulitis and impetigo), to more sever conditions (eg. osteomyelitis, septic arthritis, post-surgical and device-related infection), through to acute life-threatening invasive conditions (eg. septicaemia, infective endocarditis, meningitis and necrotising pneumonia) (7, 13). In the pre-antibiotic era, the prognosis of patients suffering from

severe *S. aureus* infection was extremely poor. One retrospective study from the late 1930s reported a mortality rate in excess of 80% in cases of septicaemia (17). The introduction of penicillin into clinical use in the early 1940s radically altered the way in which infectious diseases were managed and provided a substantial reduction in the associated mortality of staphylococcal disease (18-20). However, staphylococcal infections have remained a major public health burden and *S. aureus* is a leading cause of nosocomial disease.

Because of the extreme disparity in disease severity, clinical manifestation, and with some ethnicities disproportionately affected (21, 22), it has been a challenge to accurately quantify the global burden of staphylococcal infection. Most studies have focused on well-defined populations and only a single or small number of clinical conditions. Bacteraemia is often investigated because in the great majority of invasive *S. aureus* infections the organism can be recovered from blood (23). For this reason, monitoring the isolates responsible for *S. aureus* bacteraemia (SAB) can provide useful insight into genomic and phenotypic shifts occurring in the wider invasive disease-causing population, such as changes in the molecular epidemiology and antimicrobial resistance profiles which may impact on clinical outcome. The incidence of SAB in industrialised countries has been estimated to range between 10 to 50 per 100,000 person-years, with case fatality ranging from 15 to 25% (24-27). In Australia and New Zealand, the incidence of SAB is estimated to range from 11.5 to 35 per 100,000 population (21, 28), and 16.9 to 41 per 100,000 population (29, 30), respectively. However, without consistent methodology limited comparisons can be made between studies from different regions.

Although SAB is one of the most well-described presentations of staphylococcal disease, by far the most common clinical manifestation is SSTI, for which *S. aureus* is one of the most frequently identified aetiological agents (31-33). Data surrounding the incidence of SSTI is biased towards severe infection as medical attention is less commonly sought when the disease is more benign. However, the data that is available has suggested that the incidence of SSTI is escalating in multiple regions (34-37). Of particular relevance to the research undertaken in this thesis is the increasing incidence of SSTI in New Zealand over the last two decades; with national hospitalisation rates for staphylococcal SSTI estimated to have doubled from 81 to 140 cases per 100,000 population between 2000 and 2011 (37). In the last decade, the incidence of SSTI in the region has remained high, associated with a substantial shift in the molecular epidemiology of local circulating *S. aureus* populations and the emergence of genetic lineages that display resistance to topical antimicrobials (22, 38, 39); a phenomenon that has been investigated further in Chapter 3 of this thesis.

Staphylococcal colonisation and disease is not limited to humans and has been reported in a range of animal hosts, from companion animals to livestock, birds and more exotic mammals (40-44). The role of *S. aureus* as a zoonotic pathogen is outside the scope of this thesis and will not be discussed further. However, it does demonstrate the ability of *S. aureus* to colonise or cause infection in a range of hosts and physical settings; an attribute that is facilitated by the extensive genetic diversity of the species and the flexible nature of its genome, making *S. aureus* a highly versatile pathogen.

1.3 The Genome of *S. aureus*

The first complete staphylococcal genomes were published in the early 2000s (45-49). The examination and comparison of these original sequences, some of the earliest examples of comparative staphylococcal genomics, provided the first glimpse into the gene composition, virulence and antimicrobial resistance repertoires, and evolutionary mechanisms that form and shape *S. aureus*. As the availability of complete staphylococcal genomes has increased, genomic trends in the species have emerged. A typical *S. aureus* chromosome is approximately 2.8 mega-bases (Mb) in length with 2500-2800 predicted protein coding sequences (CDS) (50, 51). It is estimated that 75-80% of the genome, equating to ~2.3 Mb, is conserved across the species both in gene composition and organisation (50, 52). These regions are termed the core genome and are primarily composed of essential metabolic, regulatory and housekeeping genes, in addition to several genes that contribute to pathogenicity (45, 50). The remaining regions, termed the accessory or auxiliary genome, are primarily composed of mobile genetic elements (MGE), a source for acquired mechanisms of virulence and antimicrobial resistance (53-55).

Currently, there are more than 43,000 *S. aureus* whole genome sequences (WGS) deposited in public repositories (56), with 358 classified as complete genomes in the National Centre for Biotechnology Information GenBank database as of the January 2019 (Figure 1.1). Although sequencing efforts in this species have been biased (56), with an overrepresentation of isolates that display enhanced pathogenic potential, antimicrobial resistance or represent clinically significant lineages, the diverse demography of this collection has revealed the extensive genomic landscape of *S. aureus* and underscores the species' potential to evolve and adapt.

1.3.1 Lineage Classification

Staphylococcus aureus, like many pathogenic bacterial species, is genetically diverse and lineage classification is key to understanding the population's structure and evolutionary history. Multiple typing methods have been developed over the last few decades enabling the establishment of population subdivisions. Several are broadly applied to bacteria, but some are genus- or even species-specific. In both cases, these methods were designed to identify variation in genetic sequences or encoded protein products, with almost all considering only a single or small number of targets.

In *S. aureus*, some of the earliest typing methods include restriction fragment length polymorphism (RFLP) and multi-locus enzyme electrophoresis (MLEE). The prior detects sequence dissimilarity by site-specific enzymatic cutting of DNA, probe hybridisation and then comparison of fragment lengths, whereas the latter assess changes in the electrophoretic properties of enzymes (57-60). *S. aureus* specific methods include *spa* and *dru* typing, with both identifying sequence variation in known variable regions; the prior targeting the *S. aureus* protein A (*spa*) gene and the latter targeting the direct-repeat unit (*dru*) located within a MGE common to methicillin-resistant *S. aureus* (61-63). Multiple alternative methods have been developed, however the most widely used are pulse-field gel electrophoresis (PFGE) and multi-locus sequence typing (MLST). An improvement of RFLP, PFGE use site-specific enzymatic cutting to fragment DNA and a modified gel electrophoresis technique to visualise the fragmentation profiles (64). A key limitation to PFGE is that a single structural change in the bacterial genome which alters the restriction enzyme target location can alter the fragmentation profile and lead to isolate misclassification.

Developed in the 1990s, MLST provided an alternative to PFGE that was less sensitive to changes in genome structure and had improved reproducibility across independent laboratories (65). MLST classification is based on allelic variation at seven highly conserved housekeeping genes. In *S. aureus* these are *arcc*, *aro*, *glp*, *gmk*, *pta*, *tpi*, and *yqil* (66). Each variant gene sequence is assigned an allelic number, with each unique combination of alleles assigned a sequence type (ST), and STs grouped into a clonal complex (CC) based on sharing six of the seven alleles with another member. The substantial number of allelic combinations makes it highly improbable for two genetically unrelated isolates to be assigned the same ST/CC, excluding recombination or horizontal transfer events involving MLST genes.

A key limitation of MLST and all typing methods discussed here is that they lack fine-scale resolution and subsequently have reduced utility in analyses that require technology, specifically WGS, that can discriminate at the level of a single nucleotide change. That aside, MLST and variations of this method (eg. core-genome and whole genome MLST) are currently the most widely employed typing method and for species such as *S. aureus*, where vertical descent is the principal mode of evolution, they have proven to be an accurate technique from which to broadly infer evolutionary history (67-70).

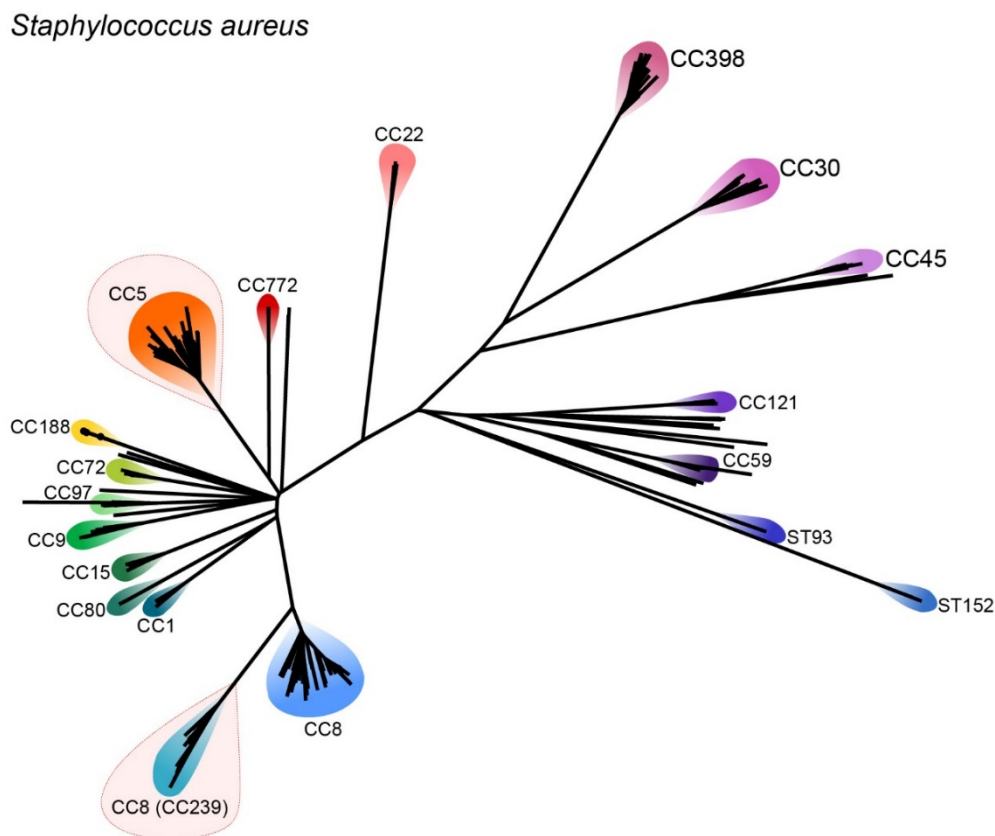


Figure 1.1 Genetic diversity of *S. aureus*. Neighbour-joining phylogenetic tree constructed from the presence/absence of 10,000 k-mers for the 358 complete *S. aureus* genome available in GenBank as of January 2019. Clonal complexes represented by more than three genomes are annotated and lineages relevant to this thesis are outlined in red.

1.4 Antimicrobial Resistance in *S. aureus*

The development of mechanism that enable microbes to survive in the presence of antimicrobial agents was a genetic inevitability given that these chemicals act to suppress growth (bacteriostatic) or are lethal to the cell (bactericidal), establishing a strong selective pressure and adaptive advantage for those microbes that have developed or acquired mechanisms to survive and replicate in the presence of these agents (71, 72). When the widespread release of antibiotics into clinical use has been compared with the earliest reports of resistance development (Figure 1.2), two phenomena are apparent: (i) pathogenic bacteria, with *S. aureus* as a prime example, have consistently developed or acquired mechanisms to combat all clinically-utilised agents, and (ii) the time between the release of the agent into clinical use and the detection of resistance is frequently short (19, 72, 73). Both phenomena can be explained by the fact that conventional antimicrobial agents are natural products (or synthetic variants); chemicals developed by microbes as defence mechanisms against other microbes (74). Therefore, the selection pressure to develop or acquire mechanisms of resistance began much earlier than the “antibiotic era”, and their clinical utility has simply intensified the need amongst bacterial populations to rapidly disseminate the genetic material encoding these pre-evolved mechanisms (75).

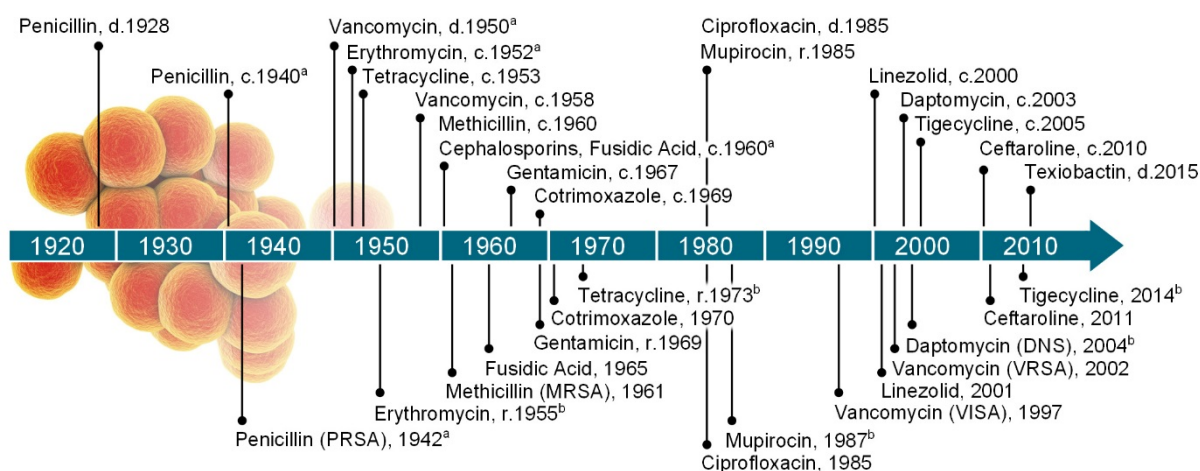


Figure 1.2 Timeline of the development and/or clinical use of antimicrobial agents and the *in vivo* emergence of resistance in *S. aureus*; modified from Figure 1 presented in (73), with data sourced from the following publications (20, 72, 74, 76-88). Year of discovery [d] and/or clinical use [c] of antimicrobial agents are presented above the timeline. Year in which resistance was first reported in *S. aureus* is presented below the timeline. ^a Indicates discrepancies between publications in reported year. ^b Indicates antimicrobials in which *in vitro* resistance was reported before *in vivo* resistance (the latter is shown).

Staphylococcus aureus has demonstrated a remarkable capacity to both develop and acquire mechanisms of antimicrobial resistance, with the identification and comprehension of these systems the focus of an extensive body of research. Early molecular biology studies focused on the effect of small chromosomal mutations in altering phenotypic susceptibility. For example, mutations that structurally change an antibiotic's cellular target to reduce binding affinity, or modulate gene expression to increase or decrease target availability (89). Despite research efforts, the spectrum of small chromosomal mutations mediating antibiotic resistance remains largely undefined because of their tremendous diversity, both in genome location and mechanism of action. Developing improved tools and techniques to explore this resistance-associated mutational landscape is the focus of current research efforts and progress is being made through merging high-throughput genomic techniques with traditional microbiological approaches (90).

Although important, chromosomal mutations alone cannot explain the rapid pace with which *S. aureus* has developed resistance to almost all classes of clinically relevant antibiotics (19, 73, 74, 91). Single- and multi-nucleotide polymorphisms (SNP, MNP) arise primarily through stochastic processes; the result of errors that occur during DNA replication and from DNA damage. Only through the accumulation of mutations over a long period of time can a genotype be achieved that is both advantageous and non-deleterious in the antimicrobial-exposed environment (52, 92). Through the later discovery of MGEs and horizontal gene transfer (HRT), it was established that antibiotic resistance in *S. aureus* was largely the result of a mobile 'resistome' (50, 52, 53, 93). The intra- and inter-cellular movement of resistance- and virulence-associated genes is enabled by a diverse range of autonomously mobile genetic entities. These elements have contributed immensely to the evolution and adaptation of staphylococcal populations and are discussed further in section 1.6, with their role in augmenting resistance investigated in Chapter 4. The remainder of section 1.4 provides an overview of the evolutionary history of antimicrobial resistance development in *S. aureus*, followed by a discussion about some of the ongoing and emerging challenges. A summary of the diversity of antimicrobial resistance mechanisms identified in *S. aureus* has been provided in Table 1.1.

1.4.1 Evolutionary History of Antibiotic Resistant *S. aureus*

The evolution and global spread of antibiotic-resistant lineages of *S. aureus* has commonly been referred to as a "wave-like" phenomena, with four well-defined waves having transpired and multiple emerging challenges on the horizon.

First Wave: Penicillin-Resistant S. aureus

The penicillins represent a group of natural and semi-synthetic chemicals that are some of the oldest and most widely used antibiotics; structurally characterised by the presence of a β -lactam ring connected to a thiazolidine ring. Although first discovered in 1928, penicillin was not extensively used as a therapeutic agent until the early 1940s (91). The penicillins are analogues for the D-alanyl[4]-D-alanine[5] (D-ala-D-ala) pentapeptide bridge precursor, the site at which cross-linking of the repeating N-acetylglucosamine-N-acetylmuramic acid units in the glycan chain occurs during peptidoglycan biosynthesis (Figure 1.3A). This process of cross-linking is termed transpeptidation and involves a group of transpeptidase enzymes known as the penicillin-binding proteins (PBPs) (94, 95). The interaction between the penicillins and the bacterial cell wall is complex (74, 96), however the chief

function of the penicillins is to inhibit the operation of the PBPs (Figure 1.3A). Direct binding to the transpeptidase active site precludes the final stage of peptidoglycan biosynthesis and results in significantly impairment to cell wall integrity (97, 98). The extent of cell lysis that follows exposure to the penicillins is accomplished through a combination of enzymatic inhibition, early termination of peptidoglycan biosynthesis, and the instigation of a deleterious cycle of cell wall synthesis and degradation, which occurs when the cell attempts to compensate for these changes (96).

Staphylococcus aureus that could survive in the presence of penicillin, termed penicillin-resistant *S. aureus* (PRSA), were first reported in the early 1940s (91, 99). Resistance was mediated by the production of an enzyme (commonly referred to as a β -lactamase or penicillinase) that could hydrolyse the β -lactam ring, rendering penicillin inactive (100, 101). In *S. aureus*, this enzyme (BlaZ) is under the regulation of a penicillin-inducible repressor BlaI and anti-repressor BlaR1, encoded by genes of the same name (91, 102). Four variants of BlaZ are recognised (types A, B, C and D), with each displaying different hydrolytic properties against the β -lactam antibiotics (103). These genes are carried as part of a transposable element in *S. aureus* (Tn552 or Tn552-like), found to be harboured by multiple plasmids (BlaZ types A, C and D) and integrated into the chromosome (BlaZ type B) (104).

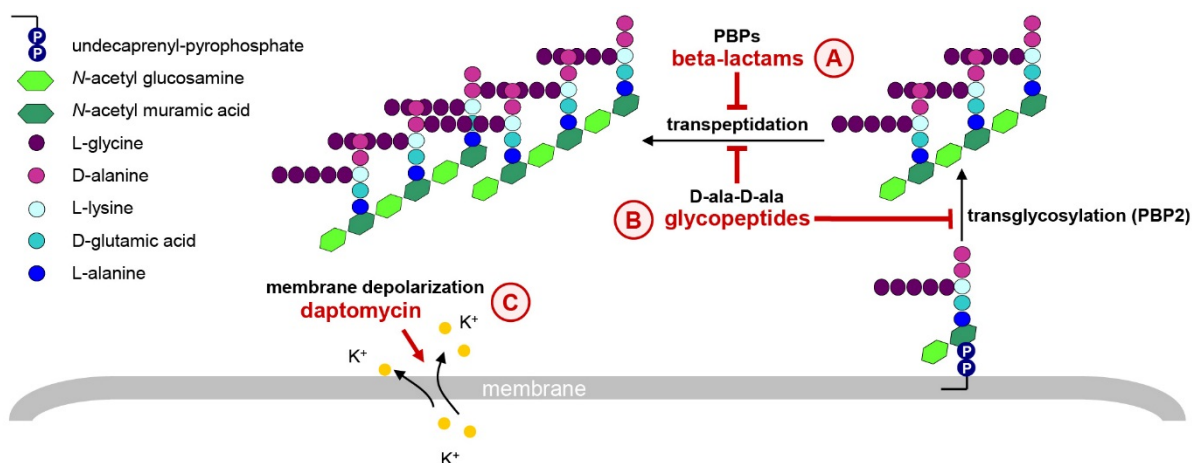


Figure 1.3 Illustration of late stage peptidoglycan biosynthesis in *S. aureus* and the mechanisms targeted by cell wall active antimicrobial agents, modified from Figure 1 presented in (94). The β -lactams [A] target transpeptidation, binding to the transpeptidase enzyme to prevent cross-linking. The glycopeptides [B] target both transpeptidation and transglycosylation, by binding to D-ala-D-ala in the pentapeptide. Daptomycin [C] when bound to calcium inserts into the cell membrane impairing structural integrity resulting in ion leakage, notably potassium efflux. Cell wall active antimicrobial agents that target earlier stages of peptidoglycan biosynthesis are not shown.

During the 1950s and 1960s, the incidence of PRSA rapidly reached pandemic levels and it is estimated that 80-90% of all contemporary *S. aureus* are penicillin resistant (19, 20, 91). Of interest, a recent report noted an increase in the prevalence of penicillin-susceptible *S. aureus* (PSSA) in a hospital in New York, United States (105). A similar phenomenon may also be occurring in Australia, as suggested by data collected as part of national surveillance of SAB (<http://agargroup.org.au/agar-surveys>). The factors contributing to this shift in penicillin susceptibility are currently undefined but are likely associated with changes in antimicrobial prescribing practices in combination with enhanced infection control measures to reduce the burden of resistant-bacterial infections.

Second Wave: Methicillin-Resistant *S. aureus*

To combat the increasing prevalence of PRSA, alterations were made to the structure of penicillin; the inclusion of larger moieties to replace the aminoadipoyl chain and sterically hinder binding of the penicillinase enzyme. These semisynthetic penicillinase-stable penicillins, termed the antistaphylococcal penicillins, include methicillin, oxacillin, flucloxacillin, (di) cloxacillin, nafcillin, and cefazolin. Methicillin was the first to be used clinically, released in 1960, but was quickly superseded by the other antistaphylococcal penicillins when it was found to be acid labile (74, 106).

Methicillin-resistant *S. aureus* (MRSA) were first identified less than a year after the clinical uptake of methicillin (107-110), with a recent temporal-phylogenetic study suggesting that the emergence of MRSA may have occurred nearly two decades prior to these early reports (75). Unlike penicillin, resistance to methicillin in *S. aureus* is mediated by the acquisition of a gene encoding a structurally modified PBP2, known as PBP2a or PBP2' (111, 112), with significantly reduced binding affinity for the antistaphylococcal penicillins. As a bifunctional transglycosylase-transpeptidase, PBP2a can substitute for the enzymatic action of all the PBPs, allowing peptidoglycan biosynthesis to continue unimpeded in the presence of high concentrations of β -lactams (113, 114). This single enzyme confers resistance to all penicillins (natural and semisynthetic), cephalosporins (except 5th generation) and carbapenems, and is used as a marker for β -lactam resistant *S. aureus*, more widely known as MRSA (19, 20). PBP2a is encoded for by the *mecA* gene, under the regulation of a repressor (*mecI*) and antirepressor (*mecR1*). All are carried by a group of structurally similar MGEs termed the staphylococcal cassette chromosome (SCC) elements (20, 54, 115, 116). Orthologs of *mecA* have been described. The *mecB*, *mecC* and *mecD* genes encode structurally altered PBP2 proteins with low affinity for β -lactams, displaying only 60-70% amino acid identity to *mecA* (117, 118). All *mec* orthologs are carried by SCC or SCC-like structures and have been identified in multiple staphylococcal and Gram-positive species and from varied hosts (117-122).

Third Wave: Healthcare-Associated MRSA

In the decades following the first reports of MRSA, isolates were increasingly recovered from hospitals around the globe (123-130). The lineages representing these early isolates for reasons that remain unclear did not persist (75), rather it was the rapid and extensive global spread of the descendants of these archaic MRSA that defined the third evolutionary wave (19). Like PRSA, the incidence of MRSA rapidly increased following early reports, with the greatest burden of infection felt during the 1980s and 1990s (19). By in large, the majority of MRSA infections prior to 1993 occurred in individuals with contact to hospitals and long-term care facilities (131, 132), and thus these early pandemic lineages were termed the healthcare-associated (HA) MRSA clones. The

evolutionary dynamics surrounding the development of HA-MRSA and their global spread were not revealed until the introduction of molecular-based typing schemes, with our knowledge of these events continuing to be refined with the use of comparative genomics.

The earliest theories about HA-MRSA were first reported in the 1990s, following the establishment of basic phylogenetic models from molecular based typing techniques, namely RFLP and MLEE. These approaches yielded two contradictory hypotheses. (i) The SCC element carrying *mecA* (SCC*mec*) was acquired by the species only once and through clonal expansion and intercontinental transmission did MRSA disseminate around the globe (58). (ii) The SCC*mec* was acquired on various occasions into genetically divergent *S. aureus*, establishing multiple HA-MRSA lineages within distinct genetic and geographic origins (57). At the time, neither study was sufficiently discriminatory to disprove the other but with the continued development and wide spread uptake of molecular typing techniques and genomics the latter hypothesis is the most plausible and by far the most widely accepted theory.

It is now known that five major lineages are responsible for the large majority of HA-MRSA infection: CC5, CC8, CC22, CC30, and CC45 (19, 67, 133). Each lineage has independently acquired SCC*mec* on multiple occasions (19, 134, 135), and demonstrated an ability to persist in healthcare environments, with all still circulating today. The enhanced survival and spread of HA-MRSA lineages has been attributed to the adaptive advantage provided in the form of an extensive and varied resistome (19), enabling HA-MRSA to exhibit resistance to multiple antibiotic classes. However, the maintenance of extended resistance in HA-MRSA appears to have come at a cost, with these lineages often displaying impaired pathogenic potential and reduced fitness compared to their susceptible counterparts (136, 137). Although well-studied due to their clinical significance, there remains gaps in our knowledge about how HA-MRSA lineages were able to thrive in the highly selective healthcare environment and the evolutionary pathways that were utilised to facilitate this; knowledge that will aid in the ongoing development of efficacious infection prevention and control measures, and in guiding antibiotic therapy and stewardship. Exploring the extended evolution of a dominant lineage of HA-MRSA in Australian hospitals to improve our understanding of these dynamics is the focus of the research presented in Chapter 2 of this thesis, and the associated publication (137).

Fourth Wave: Community-Associated MRSA

The next wave in the evolution of antimicrobial resistant *S. aureus* is defined by the emergence of the community-associated (CA) MRSA lineages, so named because they were primarily recovered from individuals without recent contact to healthcare environments. The earliest reports of CA-MRSA are from the early 1990s, where multiple MRSA cases were identified from indigenous communities in the remote Kimberly region in north-western Australia (131); the patients having had no contact with healthcare environments and subsequently no obvious source for the infection. This was proceeded by similar reports in North America and Europe (138-141). In addition to the demographic differences of patients infected with CA-MRSA, these isolates demonstrated multiple bacterial attributes that distinguished them from HA-MRSA: (i) CA-MRSA were typically resistant to a limited number of antibiotics, a finding thought to result from restricted exposure to these agents outside of the healthcare environment. (ii) While HA-MRSA is recovered predominantly from patients with predisposing risk factors for infection, CA-MRSA infection and carriage is common amongst otherwise healthy individuals, suggesting

enhanced pathogenic potential (141). Molecular and genomic analysis of CA-MRSA supports these findings, with the identification of a limited resistome but extensive virulome when compared to HA-MRSA. Further, these molecular and genomic investigations revealed that (iii) CA-MRSA represent genetically distinct lineages of *S. aureus*, having independently acquired unique variants of *SCCmec* (19, 140-142). While some of these lineages overlap with HA-MRSA, genetically distinct sub-lineages are observed to predominate in the healthcare vs community environments (19, 37, 141, 143).

Today, CA-MRSA is now far more common than HA-MRSA. In Australia, 76.9% of MRSA collected for national surveillance of SAB in 2018 were classified as belonging to CA-MRSA lineages (<http://agargroup.org.au/agar-surveys>). In New Zealand, CA-MRSA represented 84.5% of isolates collected during the 2017 national MRSA survey, a month-long snap-shot of clinical MRSA (https://surv.esr.cri.nz/antimicrobial/mrsa_annual.php). In both cases, these are an underrepresentation of the proportion of CA-MRSA as these surveillance programs by design do not capture the extent of staphylococcal infection occurring in the community unless the disease is sufficiently severe for the patient to seek medical attention.

While the terms HA-MRSA and CA-MRSA are still regularly used, the attributes considered when making these classifications have changed. Historically the distinction was an epidemiological one, based on whether the patient had recent exposure to a healthcare setting. This distinction is still made but more commonly termed “community-onset” for patients with no recent healthcare contact and with infection onset < 48 hours after hospital admission; “hospital-onset” for infections presenting > 48 hours after hospital admission; and “hospital with community onset” for patients that would be designated community-onset but have a known history of healthcare exposure; definitions that are primarily used to assess which infection control and prevention measures are required to preclude transmission (142, 144). For the most part, the terms HA-MRSA and CA-MRSA now reflect genetic differences in MRSA lineages, namely the evolutionary history of an isolate (*i.e.* whether they are a descendant of a historic HA- or CA-MRSA lineage), the number of antimicrobial classes for which the isolate is predicted to be resistant, and the variant of *SCCmec* that is carried. There are however an increasing number of exceptions to these definitions as the distinctions between the two populations becomes more ambiguous (69).

1.4.2 Ongoing Challenge of Glycopeptide and Daptomycin Resistance

Glycopeptides & Lipoglycopeptides

Vancomycin was first introduced into clinical use in 1958, having been discovered as a natural product produced by *Amycolatopsis orientalis* (previously known as *Streptomyces orientalis*), and used in the treatment of PRSA until the development of methicillin (145, 146). Following the emergence of MRSA, there was a resurgence in the use of vancomycin to treat resistant staphylococcal infections and it has remained a mainstay in the treatment of serious invasive MRSA disease. The glycopeptides (vancomycin and teicoplanin) act by binding to the D-ala-D-ala residue of the newly synthesised pentapeptide in the cell wall, inhibiting PBP2 and triggering early termination of peptidoglycan biosynthesis (Figure 1.3B). Like the β -lactams, interference with glycan chain cross-linking results in a structurally weakened cell wall, with the effects amplified as more of the agent accumulates (73). The related, semi-synthetic lipoglycopeptides (eg. dalbavancin, oritavancin and telavancin) in addition to interfering with peptidoglycan biosynthesis cause significant damage to the cell membrane and have

demonstrated potential functionality against persister cells, a transient state of metabolic dormancy (147, 148). Unlike most antimicrobial agents, the development of resistance to the glycopeptides in *S. aureus* was first reported more than three decades after its clinical use. This delay is likely the result of the complexity of the glycopeptide resistance mechanisms utilised by *S. aureus*, of which there are two well defined types.

The first type involves acquisition of the plasmid-borne transposon, Tn1546, carrying the *vanA* gene cluster. Resistance is mediated by the replacement of the target peptidoglycan precursor D-ala-D-ala with the alternative structure D-alanyl[4]-D-lactate[5] (D-ala-D-lac), for which the glycopeptides have reduced binding affinity (74, 149, 150). Termed vancomycin-resistant *S. aureus* (VRSA) because of the high-level phenotypic resistance displayed by these isolates, this mechanism is rare; the first case reported in 2003 with a further 14 cases reported since (73, 151). Carriage of Tn1546 in *S. aureus* has been associated with reduced cell fitness, plasmid instability and the encoded D-ala-D-lac has been noted to exhibit incompatibility with the native PBP2, collectively creating a highly unstable mechanism which when combined with the required co-colonisation or co-infection of the isolate with a donor vancomycin-resistant *Enterococcus*, likely accounts for the limited number of reported VRSA cases (74, 152).

Far more common and posing a greater clinical challenge is the second type of glycopeptide resistance, termed vancomycin-intermediate *S. aureus* (VISA) because of the low-level phenotypic resistance displayed by these isolates. Resistance in this case is predominantly mutation-based and largely driven by exposure to glycopeptides, such as during extended treatment for invasive staphylococcal disease. The mutational pathways leading to the VISA genotype are heterogenous but typically involve small alterations to the function of global regulatory and cell wall associated systems (145, 153-156). Phenotypic changes resulting from these mutations are commonly a thickened cell wall, increased production of D-ala-D-ala residues, decreased autolytic activity, modified surface protein profiles, and altered expression of PBPs and reduced transpeptidation; collectively facilitating the production of an abnormal cell wall structure which contributes to resistance through impaired drug diffusion, decreased drug concentrations at the division septum, and protection of the binding target (73, 145).

Development of the VISA genotype is frequently multi-step and often preceded by the emergence of low-level resistance or reduced susceptibility in a sub-population, termed heterogeneous VISA (hVISA). These sub-populations are a challenge to detect with traditional antimicrobial susceptibility testing as they commonly display a reduced growth rate, a minimal shift in *in vitro* susceptibility, and their low frequency in an otherwise susceptible population, and consequently often go undetected. In a patient, typically one with a high-bacterial-load infection (eg. endocarditis, osteomyelitis, septic arthritis, deep abscesses and device-related infections), the development of a hVISA or VISA phenotype has been associated with failure of glycopeptide therapy, persistent infection and overall poorer clinical outcomes (145, 157, 158). Multiple *in vitro* methods have been developed to detect staphylococcal (sub)populations with reduced glycopeptide susceptibility, their varied application leading to ambiguity surrounding the global prevalence of hVISA and VISA (157). Although, there does appear to be some lineages of *S. aureus*, namely the ST239 HA-MRSA (CC8) lineage that is investigated in Chapter 2, that have a genetic predisposition to developing reduced glycopeptide susceptibility (158).

Daptomycin

A commonly prescribed alternative to the glycopeptides is daptomycin, a cyclic peptide with a decanoyl fatty acid side chain, produced by *Streptomyces roseosporus* (159, 160). Daptomycin requires calcium to form a cationic complex which can oligomerise. Subsequent insertion into the cell membrane leads to structural weakening, depolarisation, and ion leakage, most notably potassium efflux, resulting in rapid cell death (Figure 1.3C). However, our understanding of the full extent of daptomycin activity remains incomplete (161, 162).

As with all clinically utilised antistaphylococcal antibiotics, resistance has developed in the form of daptomycin non-susceptible (DNS) *S. aureus* (Figure 1.2), with guidelines to define daptomycin resistance recently released by the European committee on antimicrobial susceptibility testing (http://www.eucast.org/clinical_breakpoints/). Resistance to daptomycin is mutation based, and commonly involves increased activity of the multi-peptide resistance factor protein (MprF), encoded by a gene of the same name. MprF is responsible for the lysinylation of the phosphatidylglycerol, creating positively-charged phospholipid structures, and their translocation to the outer cell membrane. Increased MprF activity leads to a general increase in positive surface charge in the bacterial cell, resulting in increased electrorepulsion of the calcium-daptomycin complex (160, 163, 164). Secondary mutations are commonly reported in global regulatory, cell membrane and cell wall associated systems, many of which are also associated with hVISA/VISA (145, 165, 166). These mutations are heterogenous and result in varied phenotypic profiles. Those commonly identified include: increased cell surface charge, altered cell membrane fluidity (the result of a shift in either fatty acid composition or phospholipid content and symmetry), altered membrane permeability, increased production of carotenoid pigments, as well as those phenotypes observed with hVISA/VISA when the mutation provides cross-resistance to the glycopeptides (160, 166-169).

Of clinical concern, exposure to daptomycin or a glycopeptide can trigger the develop of mutation(s) that provide cross-resistance to both classes of antibiotics, therefore eliminating a key treatment alternative for patients infected with such strains. The prevalence of DNS amongst clinical *S. aureus* with reduced glycopeptide susceptibility has been reported to range from 13-80%, with variation noted between the methodologies used to assess resistance (170-172). Adding further complexity to the problem are the cationic antimicrobial host defence peptides (HDP). These are biologically active molecules produced by the host that aid in immune activation, inflammation, and clearance of microbes (173, 174). As with antibiotics, bacteria have developed mechanisms to resist the action of HPDs but as there is overlap between the cellular targets of daptomycin and some HPDs, specifically the negatively charged components of the cell membrane, mutations developed in response to one can provide cross-resistance or prime the bacteria for resistance development to the other (175); a phenomenon that is not limited to *S. aureus* or daptomycin, but serves as another example of the need to develop novel antimicrobials and therapeutic approaches to tackle antimicrobial resistant pathogens (176, 177).

1.4.3 Emerging Challenge of Topical Antibiotic and Disinfectant Resistance

Skin and soft tissue infections represent the most common and varied clinical manifestation of staphylococcal disease, and colonisation with *S. aureus* is a recognised risk factor for infection (10, 15, 16). Examples of SSTIs include impetigo, cellulitis, cutaneous abscesses, ulcers, burn-associated wounds, and surgical site infections (13, 178). The management and prevention of staphylococcal SSTI can include the use of a topical antibiotic, with topical agents having multiple potential benefits over systemic agents for impetigo (179). These include: ease of application, less risk of toxic side-effects, and delivery of high drug concentrations to the infection site (178). Other measures commonly used in the prevention of staphylococcal SSTI rely heavily on biocides, often in combination with topical antibiotics, to reduce or eliminate carriage of *S. aureus* on the skin. Chlorhexidine is by far the most widely used biocide for decolonisation and disinfection due to its broad spectrum and residual activity against bacteria, fungi, and some viruses. It is also widely tolerated, with few side-effects when used as a topical agent (180-183). Unsurprisingly, the increasing use of topical antibiotics and disinfectants has driven the acquisition and development of mechanisms of resistance and/or tolerance in *S. aureus*. Of relevance to the research presented in this thesis is the topical antibiotic fusidic acid (Chapter 3) and the biocide chlorhexidine (Chapter 4).

Fusidic Acid

Fusidic acid (FA) is a steroid-like fusidane chemical produced by *Fusidium coccineum* (184), and acts on the bacterial cell to inhibit protein synthesis (185, 186). This is accomplished when FA binds to elongation factor G (EF-G), a cytoplasmic protein that functions as the sole ribosomal translocase in prokaryotes. When EF-G is bound to the ribosome, binding of FA prevents dissociation of EF-G, stalling the ribosome and preventing further elongation of the polypeptide chain (186, 187), as illustrated in Figure 1.4A. Clinical use of FA has led to the development of resistance in *S. aureus*, with two mechanisms having been described. (i) EF-G is encoded by the gene *fusA* and mutations in this gene, the most commonly described occurring in structural domain III, alter the shape of the encoded protein which reduces the binding affinity of FA (188-190) (Figure 1.4B). (ii) Involves acquisition of a FusB-type resistance determinant. Encoded by *fusB* or *fusC*, these metalloproteins act to protect EF-G by binding to a non-FA target site on the translocase, hypothesised to indirectly obstruct FA binding. Further, binding of a FusB-type protein to EF-G encourages dissociation from the ribosome and reinitiates peptide synthesis (Figure 1.4C), a process that is further augmented in the presence of bound FA (191, 192).

Since the 1960s, FA has been utilised as monotherapy and in combination with other antibiotics to treat both invasive and superficial staphylococcal disease. As a topical antibiotic, it is available as an ointment, cream, lotion or gel preparation, enabling easy application to superficial infection sites and for skin decolonisation. Resistance to FA was reported shortly after its release (Figure 1.2), and large contemporary population studies have demonstrated that the prevalence of resistance varies widely between different geographic locations (178). In New Zealand, FA is unregulated (available over the counter at local pharmacies) and following the implementation of local restrictions to mupirocin use, FA dispensation rates have steadily increased over the last two decades (39). Concurrently, an increase in the prevalence of FA resistant *S. aureus* has been observed, from 17% in 1999 to 25-28% in 2013-2014 (38), with a similar phenomenon also reported in the United Kingdom (193). There is growing interest in understanding what is occurring in New Zealand to both halt the local spread of resistance and preserve

the long-term utility of FA. The research presented in Chapter 3 and the associated publication (194), represents one aspect of our investigations surrounding this event.

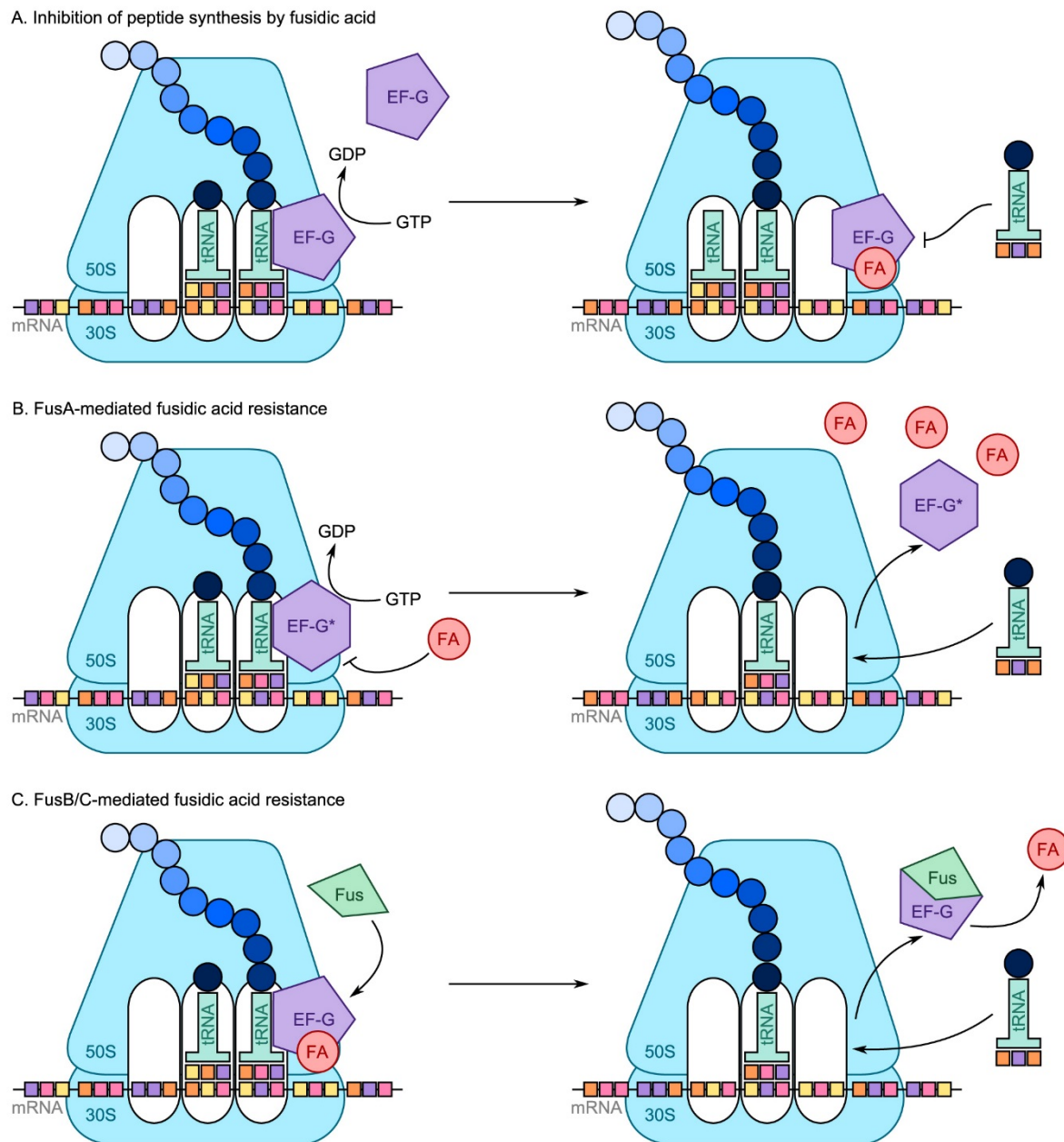


Figure 1.4 Illustration of ribosomal peptide synthesis, mode of action of fusidic acid and mechanisms of resistance, modified from Figure 5 presented in (192). (A) FA binds to EF-G preventing its dissociation and causes the ribosome to stall. Recognised mechanisms of FA resistance include: (B) development of mutation in *fusA*, which can result in the production of a structurally altered EF-G with reduced binding affinity for FA, or (C) acquisition of *fusB* or *fusC*. Both encode a fusB-type protein which can bind EF-G in an independent location to FA to prevent FA binding or promote EF-G dissociation from the ribosome when FA is bound and restart translation.

Chlorhexidine

Chlorhexidine is a cationic biguanide used widely as an antiseptic, disinfectant and in skin decolonisation, typically at concentrations of 0.5-4% in a water-soluble gluconate form (181). It acts on bacterial cells by binding to the exposed anionic components of the cell wall and cell membrane, causing formation of bridges between adjacent phospholipid heads and displacing bound cations; the effect of which is structural destabilisation and increased membrane permeability. Further, at high intra-cellular concentrations, chlorhexidine causes cytoplasmic solidification. All of which leads to cell death (195). The multiple targets and mode of action of chlorhexidine pose a significant evolutionary challenge for the development of resistance and after more than 50 years of clinical use, this remains an obstacle that *S. aureus* has yet to overcome (181). However, in response to the widespread use of chlorhexidine, *S. aureus* and many other bacterial species have acquired mechanisms that mediate tolerance, defined as a reduction in chlorhexidine susceptibility compared to natural populations (180).

In *S. aureus*, efflux is the most common mechanism for reduced chlorhexidine susceptibility (196). The genes encoding biocide-active efflux systems can be classified into two groups, the major facilitator superfamily (MFS) and small multidrug resistance (SMR) family. In both cases, the encoded proteins mediate efflux of a range of cationic and lipophilic substrates (180, 197). However, only a subset of the MFS genes encode chlorhexidine-active efflux pumps. The most commonly identified include the quaternary ammonium compound (*qac*)A and *qacB* genes, which encode peptides that differ by a single amino acid (197, 198), and are carried by a range of staphylococcal plasmids (180). An alternative mechanism, mutations in the conserved efflux pump encoded by *norA* and the associated *norB* have been associated with phenotypic shifts in chlorhexidine susceptibility, as well as other antibiotics and biocides (199-203). A current conundrum with *qac*-mediated chlorhexidine tolerance in *S. aureus* is that an increase in *qacA/B* carriage in a population inconsistently correlates with a population-level shift in susceptibility. Further, reduced chlorhexidine susceptibility cannot always be attributed to gain of, or mutation in, a known efflux system (180). These findings allude to a more complex and multi-faceted mechanism of tolerance, and unravelling the intricacy of this mechanism is one of the key aims of the research presented in Chapter 4 of this thesis, and the associated publication (204).

Table 1.1 Summary of antimicrobial resistance mechanisms in *S. aureus*

Antibiotic	Resistance mechanism	Associated MGEs
Cell wall active antibiotics		
Penicillin	PRSA: Acquisition of <i>blaZ</i> , encoding a penicillinase capable of hydrolysing the β -lactam ring (101, 205).	Tn552 & Tn552-like, commonly plasmid-borne (104)
All β-lactams (ex. 5G cephalosporins)	MRSA: Acquisition of <i>mecA</i> or <i>mecC</i> , encoding a low-affinity PBP2a enzyme (111, 112, 117).	SCC elements (54)
	BORSA: Mutation in <i>blaZ</i> resulting in hyperproduction of the penicillinase enzyme (206).	-
	MODSA: Mutation in native PBP genes resulting in altered protein structures with reduced β -lactam affinity (207).	-
5G Cephalosporins	Mutation in <i>mecA</i> , encoding altered protein structure with reduced affinity for ceftaroline and ceftobiprole (208, 209).	-
Glycopeptides	VRSA: Acquisition of <i>vanA</i> , encoding an alternative peptidoglycan precursor (D-ala-D-lac) with reduced glycopeptide affinity (73, 210).	Tn1546 (149)
	VISA & hVISA: Heterogenous mutation-based resistance. Commonly affected genes are those associated with global regulation and cell wall synthesis, including <i>walkR</i> (<i>yycFG</i>), <i>graRS</i> , <i>vraRS</i> , <i>rpoB</i> , <i>rpoC</i> , <i>msrR</i> , <i>fdh2</i> , <i>sle1</i> , <i>cmk</i> (145, 153, 154, 156).	-
Lipoglycopeptides	Mutations in genes associated with other resistance mechanisms have been identified in dalbavancin non-susceptible <i>S. aureus</i> (eg. VISA: <i>vraT</i> , <i>rpoB</i> ; MODSA: <i>pbp</i> ; Fusidic Acid: <i>fusA</i>), but none have been conclusively proven to cause lipoglycopeptide resistance (211, 212).	-
Disruption of cell membrane		
Lipopeptides (Daptomycin)	DNS: Heterogenous mutation-based resistance. Mutations typically result in electrostatic inhibition or structural alteration to cell wall and cell membrane. Multiple genes have been implicated, including <i>mprF</i> , <i>walkR</i> (<i>yycFG</i>), <i>rpoB</i> , <i>rpoC</i> , <i>dltA</i> , and <i>cls2</i> (165, 168, 213).	-
Interference with protein synthesis (targets 30S ribosomal subunit)		
Aminoglycosides	Most common mechanism is acquisition of cytoplasmic aminoglycoside modifying enzymes (AME) which alter the antibiotic to prevent binding. AMEs include the N-acetyltransferases (AAC), O-phosphotransferases (APH), and O-adenyltransferase (ANT), encoded by genes of the same name with numerous allelic variants (214). AME encoding genes of interest to the research presented in this thesis are <i>aacA(6')</i> - <i>aph(2'')</i> , <i>aph(3')</i> -IIIa, and <i>aadD</i> .	AMEs are associated with numerous MGEs and can be found both chromosomally or extra-chromosomally located (104, 214). <i>aacA(6')</i> - <i>aph(2'')</i> is carried by Tn4001 (215); <i>aph(3')</i> -IIIa by prophage SP β -like (216), and <i>aadD</i> by plasmid pUB110, occasionally integrated in SCCmec (217).

Antimicrobial Agent	Resistance mechanism	Associated MGEs
Tetracyclines	Acquisition of genes encoding tetracycline-active efflux pumps (MFS proteins encoded by <i>tetK</i> and <i>tetL</i>) or ribosomal protection proteins which preferentially bind to the target site, dislodging tetracycline (<i>tetO</i> , <i>tetM</i>) (79).	<i>tetK</i> is carried by pT181, integrated in SCCmec, <i>tetO</i> by Tn916, and <i>tetM</i> by Tn1545 (79, 104).
Glycylcyclines (Tigecycline)	Mutation-based, involving alteration of ribosomal target site or increased expression of the efflux pump encoded by <i>mepA</i> (218).	-
Interference with protein synthesis (targets 50S ribosomal subunit)		
Macrolides [M], Lincosamides [L], Streptogramins [S_A, S_B], Oxazolidinones [O], Florfenicol [Ph], and Pleuromutilins [P]	Due to the overlap in the target and/or action of these agents, acquired and mutation-based resistance mechanisms commonly provide cross-resistance. Acquisition of genes encoding ribosomal protection proteins (eg. rRNA methyltransferases which provide steric interference or proteins that induce structural changes in the drug target to reduce binding affinity), and/or antibiotic-active efflux mechanisms (219-221). For example: <i>erm</i> (MLS _B); <i>cfr</i> (PhLOPS _A), ABC-F proteins: <i>vga</i> , <i>lsa</i> , <i>sal</i> (LSA _P); <i>msr</i> (A) (MS _B) Mutation-based, resulting in alteration to the structure of ribosomal proteins, notably <i>rrl</i> [23S], <i>rplC</i> [L3], <i>rplD</i> [L4], and <i>rplV</i> [L22], can reduce binding affinity. Alteration in the expression of genes involved in rRNA modifications can also contribute to reduced susceptibility (eg. <i>rlmN</i>) (222, 223).	<i>erm</i> , <i>cfr</i> , <i>vga</i> , <i>lsa</i> , <i>sal</i> , <i>msr</i> are carried by plasmids and/or transposons (53)
Interference with protein synthesis (targets cytoplasmic proteins)		
Fusidic Acid	Mutation in <i>fusA</i> , encoding altered protein structure with reduced binding affinity. Most common mutations occur in structural domain III (190). Acquisition of <i>fusB</i> or <i>fusC</i> ; encoding FusB-type metalloproteins which bind to EF-G preventing binding of and/or dislodging bound fusidic acid (192).	- <i>fusB</i> is carried by plasmids, (224, 225), and <i>fusC</i> is carried by SCC <i>fusC</i> (193).
Mupirocin	Mutation in target isoleucyl-tRNA synthetase (<i>ileS</i>) can reduce binding affinity of mupirocin, mediating low-level phenotypic resistance (74, 226). Acquisition of <i>mupA</i> (<i>ileS</i> 2), or <i>mupB</i> , encoding mupirocin-insensitivity copies of <i>ileS</i> , mediates high-level phenotypic resistance (227).	- <i>mupA</i> is plasmid-borne. Eg. pUSA03 and pNZAK1 (104, 228).
Interference with folic acid synthesis		
Trimethoprim	Mutation in the target dihydrofolate reductase (DFR) can reduce binding affinity. Eg. <i>drfA</i> _{F98Y} (229). Acquisition of <i>drfA</i> , <i>drfG</i> , or <i>drfK</i> , encoding trimethoprim-insensitive copies of DFR, mediates high-level phenotypic resistance (230-232).	- Acquired <i>drf</i> genes are carried by transposable units, often located on plasmids (104).
Sulphonamides	Mutation in the target dihydropteroate synthetase can reduce binding affinity (233, 234).	-

Antimicrobial Agent	Resistance mechanism	Associated MGEs
Interference with DNA replication		
Fluoroquinolones	Mutations in target DNA gyrase and topoisomerase IV genes (<i>gyrA</i>, <i>gyrB</i>, and <i>parC</i>, <i>parE</i>, respectively) can reduce binding affinity. Accumulation of mutation in the quinolone resistance determining region (QRDR) augments resistance in a step-wise manner. Multiple QRDR mutations are required for clinically-relevant phenotypic resistance (235). Mutations in conserved efflux mechanisms (<i>norA</i>), or transcription regulator <i>mgrA</i> can lead to increased efflux of fluoroquinolones and high-level phenotypic resistance. (235, 236).	-
Interference with transcription		
Rifampicin	Mutation in the target β-subunit of the RNA polymerase (<i>rpoB</i>), specifically in the rifampicin resistance determining region (RRDR), can reduce binding affinity (237).	-
New Antibiotics		
Teixobactin	Teixobactin targets the lipid moiety of lipid II, interfering with cell wall synthesis. There is currently no reported teixobactin resistance in <i>S. aureus</i>, suspected to be the result of a more complex mode-of-action and conservation of the lipid II structure (76, 238).	-

1.5 Mechanisms of Virulence in *S. aureus*

Staphylococcus aureus is an adept pathogen and while its antimicrobial resistance repertoire provides an advantage in surviving therapeutic treatment, it is its extensive armamentarium of virulence factors that underpins its pathogenicity. Both the genes encoding these virulence factors (VFs) and the mechanisms by which they function are not unique to *S. aureus*, with many pathogenic and some opportunistic bacteria harbouring both gene homologues and paralogues, aiding them to overcome the same host factors. Although not investigated in this thesis, it is worth stating that the interaction between *S. aureus* and a host is not one-sided, and host-specific factors are central in moderating the onset, progression and outcome of staphylococcal disease. *S. aureus* maintains strict control over the expression of VFs through multiple regulatory systems but these are influenced by host and environmental factors, and it is when the balance between these entities is lost that *S. aureus* transforms from a coloniser to a pathogen (15, 239).

In *S. aureus*, VFs can be broadly classified into the surface-bound and secreted proteins, and then sub-classified based on the function they perform, with many VFs performing diverse tasks across different functional categories. As a broad summary these are:

Adherence to host cells. The microbial surface components recognising adhesive matrix molecules (MSCRAMMs) represent a diverse group of surface bound VFs that aid in adhesion to host cells. These include: clumping factors (*clfA*, *clfB*), fibronectin binding proteins (*fnbA*, *fnbB*), collagen binding protein (*cna*), and elastin binding protein (*ebp*) (240-242). The variable carriage of these genes and specificity of the encoded MSCRAMM proteins to host factors may predispose lineages of *S. aureus* to certain disease manifestations, or cause tissue tropism under select host and/or environmental conditions (15). Of relevance to the MRSA lineage investigated in Chapter 2 (ST239) is the newly recognised surface bound VF SasX, encoded by a gene of the same name and harboured by the bacteriophage SP β -like. SasX is hypothesised to assist in colonisation through mediating enhanced adhesion to epithelial cells and abiotic surfaces, and increased aggregation of bacterial cells. The latter characteristic may also play a role in immune evasion by interfering with phagocytosis (243).

Invasion of host cells. *S. aureus* produces a number of secreted proteins (exoenzymes) that add in cell invasion, predominantly through mediating damage of host tissues. These include: proteases (*aur*, *sspA*, *sspB*, *sspC*, *splA-F*, *ScpA*), lipases (*geh*), nucleases (*nuc*), and hyaluronidases (*hysA*) (15, 244). Further, components of the staphylococcal cell wall (*i.e.* the teichoic and lipoteichoic acids) also contribute to host cell invasion by promoting a local immune response, which itself instigates host tissue damage. Strains with a higher cell wall teichoic acid content have been shown to be more immunogenic and associated with abscess formation (245). Another group of VFs interfere with host hemostasis, such as the staphylocoagulase (*coa*) and the von Willebrand Factor binding protein (*vWbp*) which bind prothrombin leading to an abnormal coagulation cascade, supporting the adherence- and invasion-associated VFs (246).

Evasion of the host immune system. Multiple secreted and surface bound VFs specifically function to hide *S. aureus* from the host immune system. These include: the antiphagocytic capsule, of which there are multiple recognised serotypes (247); the MSCRAMM protein A (*spa*), which can bind to the Fc portion of immunoglobulin G to interfere with opsonisation (248); the chemotaxis inhibitory protein (*chp*), extracellular adhesion protein

(*map*) [a major histocompatibility class II analogue], staphylococcus complement inhibitor (*scn*) protein, and staphylokinases (*sak*), all of which contribute to disrupt the complement pathway and interfere with neutrophil recruitment to the site of infection (15, 244, 249).

Another group of VFs involved in modulating the host immune system are the toll/interleukin-1 receptor (*tir*) domain containing proteins. In *S. aureus*, *tirS* (*tirS*) has been identified to contribute to immune evasion by reducing bacterial cell recognition through toll-like receptors, significantly diminishing local inflammatory responses (250). Further, *tirS* is invariably carried by *SCCfusC*, and exposure to sub-inhibitory concentrations of FA promotes overexpression of *tirS* (47, 250).

Destruction of host defences. Another sub-group of VFs that contribute to immune evasion are those that function by directly damaging host cells, notably the cytolytic exotoxins which affect erythrocytes and leukocytes. These include the pore-forming toxins: the haemolysins (*hla*, *hly*, *hld*, *hlgA*, *hlgB*, *hlgC*), the Pantan-Valentine leukocidin ([PVL] *lukFS-PV*), and leukocidin E/D (*lukE*, *lukD*) (244, 251). While some of these exotoxins are largely conserved across the species (for example the haemolysins which are considered part of the core staphylococcal genome, although not always functional), others like PVL represent acquired mechanisms of virulence, the *lukFS* genes are mobilised by bacteriophage (249, 252). The phenol-soluble modulins (*psma*, *psmβ*) represent a further group of conserved, secreted cytolytic toxins that more broadly target eukaryotic cells, in addition to modulating immune responses and aiding in biofilm formation (253).

Proteins involved in toxin-mediated disease. A second group of staphylococcal exotoxins function as superantigens, stimulating proliferation of lymphocytes with low antigen specificity. The resulting immune response is disorganised, ineffective and destructive, and contributes to certain disease manifestations. For example, the enterotoxins (*sea*, *seb*, *sec*, *sed*, *seh*, *selk*, *sell*, *selq*) are associated with staphylococcal food poisoning, the toxic shock syndrome toxin (*tsst-1*) with the disease of the same name, and the exfoliative toxins (*eta*, *etb*) with scalded skin syndrome and bullous impetigo (249, 254).

Nutrient scavenging. Closely associated with the destructive VFs described above, are the proteins that *S. aureus* utilises to source nutrients from the damaged host tissues. Two examples of nutrient scavenging VFs, both involved in iron acquisition, are the high affinity iron binding siderophores (eg. staphyloferrin A & B [*sfa*, *sbn* operon]) which capture extracellular iron bound from host transferrins, and the iron-regulated surface determinants (*isd* locus) which work to capture heme-complexed iron following toxin-mediate destruction of erythrocytes (239, 255, 256). Complementary to scavenging, are several systems that *S. aureus* relies upon to sense, regulate and utilise host elements, such as glycose, oxygen and nitrogen (239).

Genes and mutations that contribute to persistence. Biofilm production (*icaABCD* and/or other genes encoding adhesion-associated and secreted proteins) is a common mechanism of persistence for *S. aureus*. Existence in a biofilm provides protection against the host immune system (both innate and adaptive) and antimicrobials, and poses a significant challenge to the treatment of device-related infections, with removal of the device often being the only way to achieve microbial clearance (257, 258). Another mechanism that *S. aureus* utilises to persist in the host is development of a small colony variant (SCV) phenotype. Driven by heterogenous mutations in regulatory and stress-response systems, SCV is associated with extended survival in host tissues, and increased

tolerance to antimicrobials and host defences. Further, many SCV-associated mutations have been shown to be unstable, allowing for reversion when conditions are more favourable for the bacterium (259-263).

Arginine catabolic mobile element (ACME). An acquired VF, ACME represents a genomic island carrying an *arc* gene cluster (*arcC*, *arcB*, *arcD*, *arcA*, and *argR*) which differs from the conserved native cluster, and/or an *opp* gene cluster (*oppA*, *oppB*, *oppC*, *oppD*, and *oppF*). The presence or absence of the two clusters defines the ACME types: type I contains both clusters, type II contain only the *arc* cluster, and type III contains only the *opp* cluster. ACME has been implicated in the success of a CA-MRSA ST8 lineage, widely known as USA300, although its contribution to disease severity is debated (264, 265). The presence of the element has been associated with enhancing colonisation capacity in *in vitro* disease models (266). Further, carriage of a secondary arginine deaminase pathway is theorised to augment the survival of *S. aureus* in acidic environments (267, 268).

1.5.1 Regulating Virulence: The Accessory Gene Regulator

Central to the pathogenicity of *S. aureus* are the systems which strictly regulate expression of the various VFs. The major global virulence regulator system in *S. aureus* is the accessory gene regulator (*agr*), a quorum sensing, self-regulating circuit that was first described in 1986 (269). The basic architecture of the *agr* system is two divergently transcribed operons (Figure 1.5). This first operon, P2, encodes a transcriptional regulator (RNAII) and four proteins (AgrACDB), with the second operon, P3, encoding a second transcriptional regulator (RNAIII) that is co-transcribed with δ -haemolysin (270). During the early stages of growth, low level transcription occurs through the P2 operon. An autoinducing peptide precursor (AgrD) is produced then both modified and transported across the cell membrane by a membrane-bound peptidase (AgrB). As *S. aureus* replicates, the mature AIP compound accumulates in the extracellular milieu and when a critical cell density is reached (corresponding with the late stages of exponential growth) it is sensed by a membrane-bound histidine kinase (AgrC). Recognition of AIP instigates intracellular phosphorylation of AgrA, which can then function as a transcription regulator of the P2 and P3 operons; the first upregulates transcription of this circuit and the second instigates transcription of RNAIII. RNAIII subsequently acts to modulate expression of VFs, downregulating those associated with host cell adhesion and upregulating those associated with cell invasion and immune evasion (271, 272).

As a global regulator of VF expression, the activity of this system is commonly measured as an indication of the virulence potential of an isolate. This can be accomplished directly through measuring transcription levels of RNAIII (165), or indirectly by assessing the production of the co-transcribed δ -haemolysin (Figure 1.5) (273). Both techniques have been employed as part of the work presented in Chapter 2 and although informative have some significant limitations. The first being that the *agr* system is not the sole regulator of virulence in *S. aureus*. Further, a reduction in virulence potential does not necessarily equate to a reduction in the bacterium's ability to cause disease. In fact, it is becoming increasingly recognised that there are certain host environments where *agr*-dysfunction may be advantageous, such as during persistent infection when it is preferable for a bacterium to be less immunogenic (274-276). The *agr* system is additionally influenced by exposure to certain antibiotics (277), however our understanding of the effect that antimicrobial therapy has on influencing *agr* activity during infection or importantly how this could be manipulated to improve therapy remains incomplete.

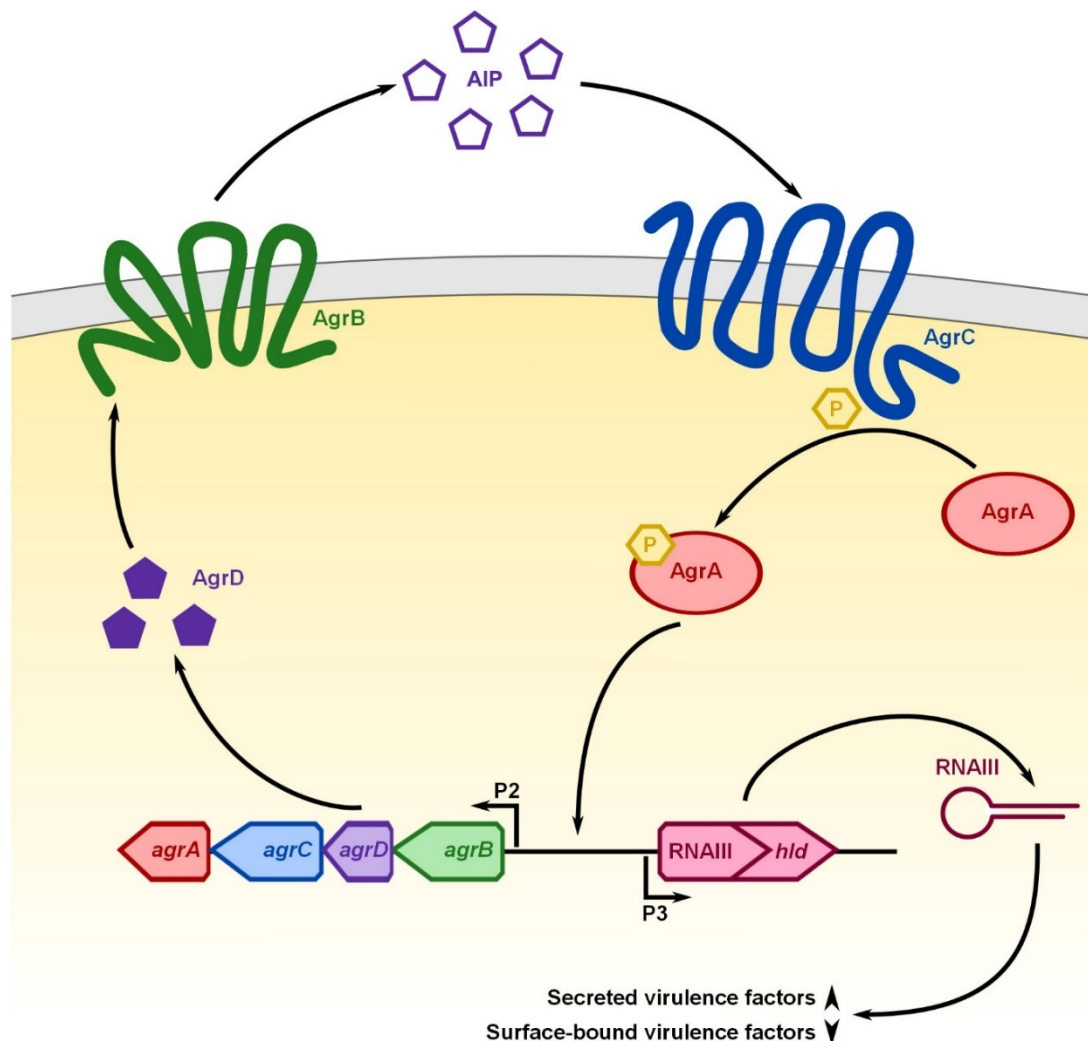


Figure 1.5 The accessory gene regulator, modified from Figure 1 presented in (278). Illustrated is a basic schematic for the *agr* system in *S. aureus*. Composed of two divergently transcribed operons, an autoinducing peptide precursor (AgrD) is produced, modified, and transported outside of the cell by a membrane-bound peptidase (AgrB). The mature autoinducing peptide (AIP) accumulates until a critic cell density is reached, which is sensed by a membrane-bound histidine kinase (AgrC). This triggers phosphorylation of the transcription factor AgrA, which acts to upregulate the *agr* system through the P2 and P3 promoters. The latter leads to production of the transcription factor RNAIII and the co-transcribed δ -haemolysin (hld). RNAIII acts to increase expression of secreted virulence factors and decrease expression of surface-bound virulence factors. Phosphorylated AgrA can also modulate the expression of non-RNAIII regulated virulence factors (not shown).

1.6 Mobile Genetic Elements in *S. aureus*

Mobile genetic elements have contributed immensely to the evolution of bacteria. These autonomously mobile entities facilitate the transfer of genetic material between two cells, referred to as horizontal gene transfer (HGT). However, they can also move within a cell and both intra- and inter-cellular movement of MGEs can influence genome structure and gene regulation (279). As a group, increasingly denoted as the ‘mobilome’, they are a key evolutionary mechanism utilised by bacteria to rapidly adapt to new and changing environments. In *S. aureus*, multiple types of MGEs have been implicated in the spread of resistance- or virulence-encoding genes, or in directly influencing these phenotypes. These include plasmids, bacteriophage, transposons, insertion sequences, pathogenicity islands, and the SCC, and have been reviewed in depth in the following publications (53, 54, 280). This section will focus on the MGEs that are of particular relevance to the work presented in Chapters 3 & 4.

1.6.1 Staphylococcal Cassette Chromosome

The Staphylococcal Cassette Chromosome (SCC) represents a group of structurally similar, large genomic islands (30 – 70kb), which are most widely known for their role in the dissemination of the *mec* resistance loci, mediating β -lactam resistance; the association so much so that the SCCs are broadly classified into those that carry *mec* (SCC*mec*) and those that don't (non-SCC*mec*). Relevant features of the SCCs include:

- **Site-specific cassette chromosome recombinases (ccr):** these proteins facilitate the excision and integration of the SCC at the 3' end of *rmlH*, previously called *orfX* (281, 282). Three *ccr* types have been identified (*ccrA*, *ccrB*, and *ccrC*), and share <50% nucleotide homology with one another. In the case of *ccrA* and *ccrB*, both have multiple described allotypes (with 50-85% nucleotide homology) and are carried together, in various allotype combination, defined as *ccr* gene complexes. A subset of non-SCC*mec* elements without function Ccrs have also been identified, and are termed ψ SCC elements (53).
- **β -lactam resistance loci:** the SCC*mec* elements are differentiated from all other SCC elements based on carriage of a *mec* gene (*mecA* or *mecC*) mediating resistance to the β -lactam antibiotics (115, 117, 119, 121). In the case of *mecA*, multiple allotypes have been described, and this gene in combination with *mecI* and *mecR1* (regulatory sequences), and the hyper-variable region and a copy of IS431 downstream of *mecA*, define the *mec* gene complex.
- **J regions:** the sequences surrounding the *ccr* and *mec* gene complexes are non-essential and highly variable, defined as the joining ‘J’ regions. In the SCC*mec* elements, these are defined as: J1, from the right chromosomal junction to the *ccr* gene complex; J2, between the *ccr* and *mec* gene complexes; J3, from the *mec* gene complex to the left chromosomal junction. These regions serve as the location for antimicrobial resistance (excluding *mecA* and *mecC*) and virulence associated genes carried by some subtypes, with the following reviews providing a comprehensive summary of these genes (53, 54).

Classification of the SCC*mec* elements is determined by the combination of *ccr* type (if *ccrA* and *ccrB* the *ccr* gene complex is used) and the *mec* gene complex, to define type, with variation in the J regions used to define sub-type. For non-SCC*mec* elements, the nomenclature is more varied and may reflect the *ccr* genes or another gene of interest that is carried by the structure, alternatively the strain from which the SCC was first identified. As of start of 2019, there are 13 officially recognised SCC*mec* types carried by *S. aureus*. Each SCC*mec* type has

multiple described subtypes and with the increasing utilisation of WGS more are being identified (53, 54, 283, 284).

Composite and chimeric SCC structures. While SCC elements are commonly found in isolation, several hybrid structures have been described (Figure 1.6). These represent the amalgamation of two or more SCC elements, termed ‘composite’ when the *ccr* gene complexes of both individual SCCs are maintained in the hybrid structure, including non-functional *ccr*, (285, 286), or ‘chimeric’ when only one of the two complexes is kept (193). Identification of these SCCs requires careful examination of the direct repeat (DR) sequences flanking the SCCs to differentiate tandem SCC integration events from true hybrid SCC structures (283).

SCC elements and *fusC*. With the exception of *mecA* and *mecC*, there is only one other antibiotic resistance gene that is carried exclusively by SCC elements: *fusC*, mediating resistance to fusidic acid. This gene is carried by a non-SCC*mec* structure named SCC₄₇₆, named for *S. aureus* strain MSSA476 in which it was identified (47). It is also referred to as SCC*fusC*. SCC*fusC* is commonly carried by MSSA lineages, notably CC1, in regions with high population use of fusidic acid (39, 193). SCC*fusC* has also been identified to form one half of multiple composite and chimeric SCC*mec* structures, which are carried by diverse MRSA lineages. For example: CC1, CC5, CC8, and CC45 (193, 194, 228) (Figure 1.6). Investigation of these elements, in particular understanding their structure, prevalence, and evolutionary history in emerging FA-resistant *S. aureus* lineages is a key aim of the work presented in Chapter 3.

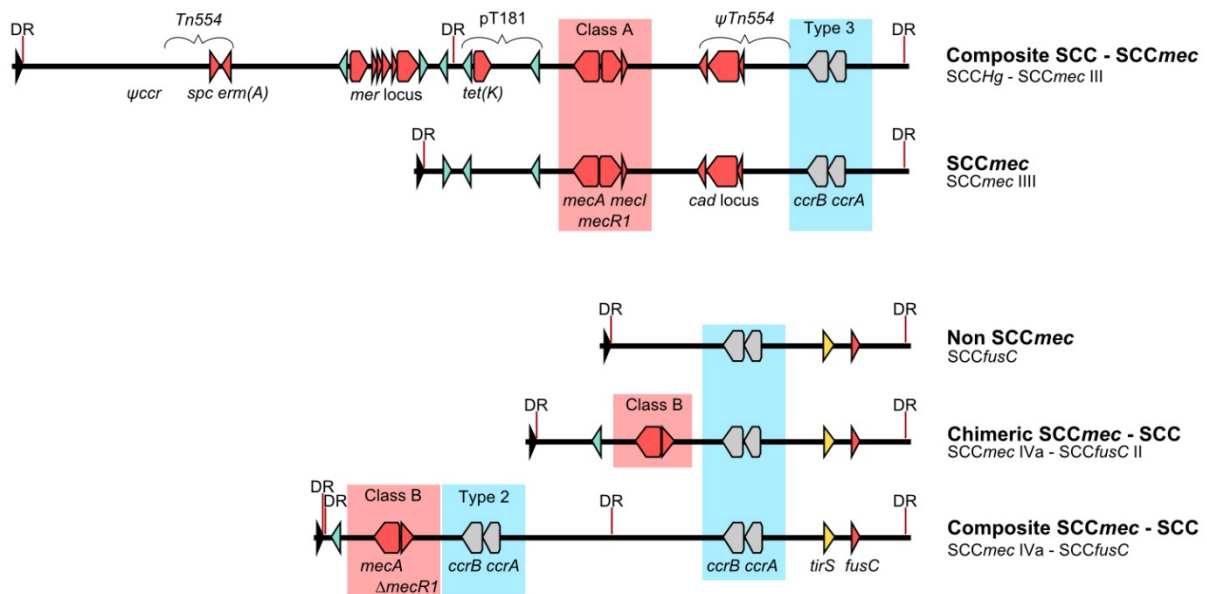


Figure 1.6 Structure of staphylococcal cassette chromosome elements. Illustrated are examples of SCC elements identified in the *S. aureus* lineages investigated in this thesis. (A) The ST239 MRSA population investigated in Chapters 2 & 4 exclusively carry a type III SCC*mec* element, either in isolation or in a composite structure with SCCHg. (B) The fusidic acid resistance *S. aureus* lineages investigated in Chapter 3 carry a *fusC*-harbouring SCC*fusC* (SCC₄₇₆) element, either in isolation (in ST1 MSSA), as a chimeric structure with a type IV SCC*mec* element in ST5 MRSA, or a composite structure with a type IV SCC*mec* element in ST1 MRSA.

1.6.2 pSK1 Plasmid Family

Plasmids are circular or linear elements of extrachromosomal DNA that can replicate independently of the chromosome and serve as vehicles for the dissemination of numerous antimicrobial resistance and virulence associated genes. In *S. aureus*, plasmids are classified into three groups based on their structure and functional characteristics (53, 104, 287). Class I: small rolling-circle plasmids. As indicated by their name, these plasmids are small (1-5kb), high copy number (15-50 copies per cell) elements that replicate using an asymmetric rolling-circle mechanism, and when associated with antimicrobial resistance, typically encode only a single mechanism. Class II: intermediate-sized, multidrug resistant plasmids, of which most can be further classified into (i) the pSK1 family of plasmids, or (ii) the β -lactamase-heavy metal resistance plasmids. All plasmids in this group undertake theta-replication, are generally low copy number, and carry multiple antimicrobial resistance determinants, as well as insertion sequence (IS) and transposon (Tn) elements. Class III: conjugative multidrug resistant plasmids. These maintain similar characteristics to the Class II plasmids but additionally encode the necessary machinery to facilitate transfer between cells by conjugation; machinery that can also mobilise other compatible non-conjugative plasmids. Of relevance to this thesis is the pSK1 family of plasmids.

Structure of the prototypic pSK1 plasmid. The sequence of the prototypic pSK1 plasmid was first published in 2010 (288), and structurally it can be divided into four regions, as illustrated in Supplementary Figure 2 in (204). Region 1 represents the plasmid backbone and includes genes encoding the replication machinery as well as those associated with the cell envelope, membrane transport, iron acquisition and a putative toxin-antitoxin system (288). This region is largely conserved amongst the pSK1 family plasmids, with the exception of integrated pSK1 plasmid variants in which the replication machinery is commonly deleted (204, 288). Region 2 represents a conjugative Tn, termed Tn4001, which carries *aac(6')-aph(2'')*, mediating aminoglycoside resistance (215, 289). Region 3 represents a second conjugative Tn, termed Tn4003, which carries a copy of *dfrA* encoding a trimethoprim insensitive DFR (230, 288, 290). Region 4, located between the two Tn harbours *qacAR*, mediating tolerance to cationic biocides (198, 291). Other pSK1 family plasmids have been found to variably contain regions 2, 3 and 4, some having acquired additional antibiotic resistance genes. For example, Tn552 which contains the penicillinase encoding *blaZ* (288).

Insertion sequences of pSK1. IS are small self-mobilising proteins that can move independently or facilitate the concurrent movement of other genes (termed a translocatable unit). They represent a diverse group of transposases (tnp) that are classified based on the protein's active site motif, which also designates the mechanism of excision/integration. The most common is DDE (aspartic acid, aspartic acid, glutamic acid) which facilitates as a cut-and-paste or copy-and-paste mechanism. From a structural perspective, the IS CDS encodes the tnp and is flanked by non-coding regions (NCR) with terminal inverted repeats (TIR), which serve as binding sites for the tnp, and are further bound within direct repeats (DR), which serve as the target sequence for integration. Two IS elements are commonly identified in pSK1 family plasmids, IS256 and IS257 (204, 288). IS256 (1324 bp) is composed of a 1173 bp CDS (encoding a DDE tnp with a copy-and-paste mechanism) with 17/26 bp terminal IRs and 8 bp DRs (292). IS257 (790 bp), belonging to the IS6-family, is composed of a 675 bp CDS (encoding a DDE tnp with a copy-and-paste mechanism) with 18/20 bp terminal IRs and 8-9 bp DRs (293). If integrated in the correct location upstream of a gene, the sequence of IS257 can also function as an alternative promoter and modulate the expression of the downstream gene, as has been observed with *dfrA* in Tn4003 (279). These

particular IS are largely responsible for changes in the configuration of the pSK1 family plasmids (204), and facilitate the mobilisation of the two composite transposons.

Composite transposons of pSK1. The historic definition of a Tn is a self-mobilising, small gene cluster with an encoded tnp (distinct from the IS-tnp and with a separate classification system) and at least one passenger gene. However, the term Tn and IS are becoming increasingly confounded, with the identification of Tn without passenger genes and IS-mobilising translocatable units, having been assigned contrary IS and Tn designations, respectively (54). The pSK1 plasmid family variable carry three Tns, Tn4001, Tn4003 and Tn552 (288), of which the first two are relevant to the work presented in Chapter 4. Tn4001 represents a 4.6 kb composite Tn, carrying *aacA(6')*-*aph(2'')* flanked by inverted copies of IS256 (215, 292). Tn4001 is capable of independent mobility and in addition to pSK1, can be found integrated into the staphylococcal chromosome at various locations. Tn4003 represents a 3.4 kb composite Tn, carrying three passenger genes including the trimethoprim insensitive *dfrA*, flanked by direct copies of IS257 (230). In pSK1, Tn4003 forms part of a larger 4.7 kb pSK639-like structure, named as it represents a co-integrate remnant of a pSK639-like plasmid (288, 290).

1.7 Research Objectives

The primary aim of the work presented in this thesis was to investigate different examples of evolution in *S. aureus* to strengthen our understanding of the dynamics surrounding staphylococcal population adaptation, with a specific focus on antibiotic-resistant lineages of *S. aureus* that circulate in Australia and New Zealand. This work endeavoured to answer questions concerning what factors contribute to the emergence and persistence of staphylococcal populations in distinct environments; what molecular tools and mechanisms were utilised to gain an adaptive advantage; and what were the consequences of these population level changes.

The continued development of WGS technology and the substantial improvements to bioinformatic software to enable comparative genomic analyses of ever-growing bacterial collections has made it possible to investigate the evolutionary dynamics surrounding bacterial populations. Nowhere else has this technology been so prosperously applied than in studying the emergence, evolution and spread of pathogenic bacteria, in particular those that have developed resistance to antimicrobials agents (41, 294, 295). However, genomic investigation in isolation can only provide hypotheses about the evolution and adaptation of bacterial populations. Greater insight is gained when this technology is paired with detailed phenotypic profiling, as this permits the exploration of the consequences of genotypic change and can reveal what characteristics are providing an adaptive advantage in a given niche.

Three examples of staphylococcal population evolution are presented in this thesis. The publication and additional data presented in Chapter 2 represent work conducted to gain a greater understanding of the ongoing evolution of the HA-MRSA lineage ST239 in Australia, one that has persisted in the region for more than four decades. Conversely, the publication and additional data presented in Chapter 3 represent work investigating the rapid emergence of a fusidic acid resistance ST5 MRSA lineage in the New Zealand community; a lineage which has now established itself as the predominate cause of MRSA infection in both the community and healthcare environments. The publication presented in Chapter 4 represents an exploration of the evolutionary dynamics surrounding the multidrug resistant pSK1 plasmid family and how changes in the configuration of this plasmid have augmented antibiotic resistance and disinfectant tolerance in the Australian ST239 MRSA population. Lastly, presented in Chapter 5 is a discussion about this work and how it has contributed to our understanding about the evolutionary dynamics surrounding successful MRSA lineages in Australia and New Zealand.

Chapter 2

**Convergent adaptation in the dominant global hospital clone
ST239 of methicillin-resistant *Staphylococcus aureus***

2.1 Introduction

One of the most successful lineages of antibiotic-resistant *S. aureus* is ST239; a multidrug resistant pandemic clone of MRSA responsible for numerous nosocomial infections since it was first reported in the late 1970s (296-298). Prior to this work only a limited number of comparative genomic studies have investigated ST239 MRSA using global isolate collections, each revealing a different aspect of this clone's evolution. (i) ST239 is a hybrid MRSA clone, formed from a large recombination event involving two major MRSA lineages, CC8 and CC30 (93). It was estimated to have emerged during the 1960s around the same time that MRSA was first identified, with potential origins in Australia or the United States (299, 300). As global ST239 collections have grown, this hypothesised geographic origin has been re-evaluated with recent data now suggesting a potential European heritage for the lineage (301), although this is still debated. (ii) The population structure of ST239 is dominated by extensive local expansion with limited intercontinental transmission. Although few, these intercontinental transmission events have successfully seeded ST239 into new regions where the population has subsequently spread, frequently displacing local HA-MRSA lineages (300-303). (iii) Although recombination has not played a major role in the evolution of ST239, since that initial event from which it was formed, there are distinct differences in recombination rates between geographic sub-clones (301). Collectively, these studies support the idea that ST239 MRSA is composed of multiple (often geographically unique) clones, each evolving independently to one another. A key limitation of this work has been the bias towards strains representing only Europe, South America and Asia; a common issue in population genomic studies because of isolate availability. Australia has only ever been represented by a single isolate, ANS46, which was recovered in Melbourne in 1982 (298, 300). Australia represents a geographically distinct region, one which is both a potential origin for this pandemic clone but also a region in which ST239 has thrived as the dominant HA-MRSA lineage for nearly four decades (<http://agargroup.org.au/agar-surveys>). Prior to this study, there had been limited data available about the local population structure of ST239 in Australia, and what evolutionary and genetic mechanisms underpinned its successful persistence in the region. Further, the phenotypic evolution of this clone had not previously been explored, specifically within the context of paired genomic data; data which could provide insight into how MRSA clones evolve and adapt within healthcare environments.

2.1.1 Project Aims

The primary aim of the work presented in this chapter was to use population and functional genomics to explore the evolution of ST239 MRSA in Australian hospitals by (i) constructing a phylogenetic model for the local population structure of ST239, and to investigate this model within the context of a global dataset to understand the clone's long-term persistence in the region; (ii) explore temporal changes in clinically relevant phenotypes, specifically antimicrobial resistance, virulence and replicative fitness, in order to understand the phenotypic adaptation of ST239 to the Australian healthcare environment, and (iii) combine these genomic and phenotypic datasets to gain deeper insight into the evolutionary and genetic mechanisms contributing to the success of this multidrug resistant lineage.

2.1.2 Chapter Overview

Presented in this chapter are the results of experiments conducted to address the three objectives described above, all of which have been answered in the publication presented in section 2.2. The supplementary methods, figures and tables from this publication are presented in section 2.2, with the supplementary dataset provided in Appendix 1. In addition to this publication, an extended genomic analysis is presented in section 2.3. In the time since this publication, WGS data for more than 400 ST239 isolates has become available. The extended analysis presented utilised this large international dataset to revisit some of the questions posed in the initial publication and provide a current picture of ST239 MRSA in Australia. A final summary of this chapter is presented in section 2.4. The implications and limitations of this work within the wider context of this thesis are discussed in Chapter 5.

2.2 Results

2.2.1 Publication: Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*

Baines, S. L., K. E. Holt, M. B. Schultz, T. Seemann, B. O. Howden, S. O. Jensen, S. J. van Hal, G. W. Coombs, N. Firth, D. R. Powell, T. P. Stinear and B. P. Howden (2015). MBio. DOI: 10.1128/mBio.00080-15.

Convergent Adaptation in the Dominant Global Hospital Clone ST239 of Methicillin-Resistant *Staphylococcus aureus*

Sarah L. Baines,^{a,b} Kathryn E. Holt,^c Mark B. Schultz,^c Torsten Seemann,^{d,e} Brian O. Howden,^a Slade O. Jensen,^f Sebastiaan J. van Hal,^{f,g} Geoffrey W. Coombs,^{h,i} Neville Firth,^j David R. Powell,^{d,e} Timothy P. Stinear,^{a,k} Benjamin P. Howden^{a,b,k}

Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology & Immunology, The University of Melbourne at The Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia^a; Microbiology and Infectious Diseases Departments, Austin Health, Melbourne, Victoria, Australia^b; Department of Biochemistry and Molecular Biology, Bio 21, Molecular Science and Biotechnology Institute, University of Melbourne, Melbourne, Victoria, Australia^c; Victorian Bioinformatics Consortium, Monash University, Melbourne, Victoria, Australia^d; Life Sciences Computation Centre, Victorian Life Sciences Computation Initiative, Melbourne, Victoria, Australia^e; Microbiology and Infectious Diseases, School of Medicine, Ingham Institute for Applied Medical Research, University of Western Sydney, Sydney, New South Wales, Australia^f; Department of Microbiology and Infectious Diseases, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia^g; Australian Collaborating Centre for *Enterococcus* and *Staphylococcus* Species (ACCESS) Typing and Research, School of Biomedical Sciences, Curtin University, Perth, Western Australia, Australia^h; Department of Microbiology and Infectious Diseases, PathWest Laboratory Medicine-WA, Royal Perth Hospital, Perth, Western Australia, Australiaⁱ; School of Biological Sciences, University of Sydney, Sydney, New South Wales, Australia^j; Department of Microbiology, Monash University, Melbourne, Victoria, Australia^k

T.P.S. and B.P.H. contributed equally to this article.

ABSTRACT Infections caused by highly successful clones of hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) are a major public health burden. The globally dominant sequence type 239 (ST239) HA-MRSA clone has persisted in the health care setting for decades, but the basis of its success has not been identified. Taking a collection of 123 ST239 isolates spanning 32 years, we have used population-based functional genomics to investigate the evolution of this highly persistent and successful clone. Phylogenetic reconstruction and population modeling uncovered a previously unrecognized distinct clade of ST239 that was introduced into Australia from Asia and has perpetuated the epidemic in this region. Functional analysis demonstrated attenuated virulence and enhanced resistance to last-line antimicrobials, the result of two different phenomena, adaptive evolution within the original Australian ST239 clade and the introduction of a new clade displaying shifts in both phenotypes. The genetic diversity between the clades allowed us to employ genome-wide association testing and identify mutations in other essential regulatory systems, including *walkR*, that significantly associate with and may explain these key phenotypes. The phenotypic convergence of two independently evolving ST239 clades highlights the very strong selective pressures acting on HA-MRSA, showing that hospital environments have favored the accumulation of mutations in essential MRSA genes that increase resistance to antimicrobials, attenuate virulence, and promote persistence in the health care environment. Combinations of comparative genomics and careful phenotypic measurements of longitudinal collections of clinical isolates are giving us the knowledge to intelligently address the impact of current and future antibiotic usage policies and practices on hospital pathogens globally.

IMPORTANCE Methicillin-resistant *Staphylococcus aureus* (MRSA) is responsible for innumerable drug-resistant health care-associated infections globally. This study, the first to investigate the evolutionary response of hospital-associated MRSA (HA-MRSA) over many decades, demonstrates how MRSA can persist in a region through the reintroduction of a previously unrecognized distinct clade. This study also demonstrates the crucial adaptive responses of HA-MRSA to the highly selective environment of the health care system, the evolution of MRSA isolates to even higher levels of antibiotic resistance at the cost of attenuated virulence. However, *in vivo* persistence is maintained, resulting in a clone of HA-MRSA able to resist almost all antimicrobial agents and still cause invasive disease in the heavily compromised hosts found in modern health care settings.

Received 15 January 2015 Accepted 28 January 2015 Published 3 March 2015

Citation Baines SL, Holt KE, Schultz MB, Seemann T, Howden BO, Jensen SO, van Hal SJ, Coombs GW, Firth N, Powell DR, Stinear TP, Howden BP. 2015. Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*. mBio 6(2):e00080-15. doi:10.1128/mBio.00080-15.

Editor Peter Gilligan, UNC Health Care System **Invited Editor** Vance G Fowler, Duke University Medical Center

Copyright © 2015 Baines et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to Benjamin Howden, bhowden@unimelb.edu.au.

Staphylococcus aureus is one of the most common and problematic causes of infectious disease in humans. The emergence and spread of epidemic drug-resistant *S. aureus* clones have often been referred to as a wave-like phenomenon, with the first wave representing the development of penicillin-resistant *S. aureus*, the second wave representing that of methicillin-resistant *S. aureus* (MRSA), and the third and fourth waves representing the es-

tablishment of the highly successful hospital-associated MRSA (HA-MRSA) lineages that are encountered today and the appearance of genetically distinct MRSA clones in the community, respectively (1).

Genomic studies are providing new insights into the population structure of *S. aureus*, MRSA in particular, and have shown that the majority of HA-MRSA infections are the result of a small

number of dominant, typically multidrug-resistant, clones that rarely cause disease outside the hospital setting (2, 3). One of the most successful and persistent clones is the globally dispersed, multidrug-resistant multilocus sequence type 239 (ST239), a naturally occurring hybrid having originated from a large recombination event involving descendants of the two major staphylococcal lineages, clonal complexes 8 and 30 (4). The earliest reports of MRSA ST239 date back to the late 1970s in Australia (5), the United Kingdom (6), and the United States (7). However, multidrug-resistant MRSA belonging to phage group III and possibly representing ST239 has been isolated in Australia since the 1960s (8). From the 1980s, ST239 had been reported as the causative agent of numerous epidemics in health care institutions worldwide; the number of regional names given to the ST239 clone attests to its success. These include, but are not limited to, the Brazilian, British epidemic (EMRSA-1, -4, -7, -9 and -11), Canadian epidemic (CMRSA-3), Hungarian, Portuguese, Nanjing/Taipei, Vienna, and Eastern Australian epidemic (EMRSA-Aus-2 and -Aus-3) clones.

Recent studies have explored the global population structure (9–11) and phylogeography (12) of ST239, as well as the role of recombination and the contribution of other mobile genetic elements to the spread and survival of this clone (11, 13, 14). However, the mechanisms underlying the successful adaptation and survival of ST239 MRSA within the health care environment, in particular, the phenotypic evolution of this clone, are poorly understood. There is a high prevalence of reduced glycopeptide susceptibility among HA-MRSA isolates in Australia (15), where ST239 has been a dominant MRSA clone for more than 30 years (16). Additionally, a recent study of *S. aureus* bacteremia isolates from Australia demonstrated higher vancomycin MICs for ST239 isolates than for other clones (17), suggesting that the ST239 clone may be adapted to enhanced antimicrobial resistance.

Using both genomic and integrated phenotypic approaches, we have explored the evolutionary dynamics of ST239 MRSA in Australia, revealing that two clades of this highly successful clone have evolved independently in our region, both contributing to enhanced antibiotic resistance and attenuated virulence driven by mutations in highly conserved essential regulatory genes. This work provides an in-depth insight into the persistence and evolution of HA-MRSA and highlights a bacterial population level response to the long-term and widespread use of antibiotics that is driving even higher levels of resistance.

RESULTS AND DISCUSSION

Global and regional phylogeny of ST239 MRSA. To recover the population structure of Australian ST239 MRSA, as well as explore the evolution of this sequence type within a global phylogeny, we collected 87 ST239 isolates from health care institutions in Australia and New Zealand between 1980 and 2012. The majority of these isolates ($n = 69$) were recovered from cases of bacteremia or tissue infection. We supplemented this collection with the publicly available whole-genome sequence reads of a further 35 international ST239 isolates from the collection of Harris et al. (10). These additional isolates were recovered from Asia, Europe, South America, and North America between 1993 and 2007 and represented the phylogenetic diversity observed in their population model. Relevant isolate demographics and a summary of the sequencing results are available in Data Set S1 in the supplemental material.

Alignment of whole-genome sequence reads for all 123 isolates with reference ST239 isolate TW20 identified 2,716 core genome single nucleotide polymorphisms (SNPs). A maximum clade credibility (MCC) tree generated from these core genome SNPs is presented in Fig. 1A. The population structure proposed by the MCC tree is robust because of the strong branch support and the distinct grouping of isolates with shared geographic source creating regional clades, a finding consistent with previous ST239 phylogenetic models. These clades reflect the capacity of ST239 to undergo extensive local expansion but limited intercontinental transmission (9–11).

The Australian ST239 isolates followed a similar pattern, grouping into two distinct clades on the MCC tree (annotated as Clades 1 and 2). The separation of these clades indicates the individual evolutionary paths along which they have developed, while the distance between them suggests that their core genomes are as different from each other as those of other continental clades. This reveals that ST239 MRSA has undergone concurrent but independent spread twice within Australia. Clade 1 isolates represented all of the regions sampled and 27 of the 32 years in the temporal span of the collection (1980 to 2007). Conversely, Clade 2 isolates were considerably less diverse; all were recovered in the Australian state of Victoria, the earliest of which was isolated in 2001. The difference in the demographic diversity of these two clades is consistent with Clade 1 being the original Australian ST239 clade and Clade 2 representing a more recent, previously unrecognized, reintroduction of ST239 into the Australian region. The more recent common ancestor shared by Clade 2 and the predominately Asian isolates, combined with the high posterior probability of major nodes connecting these groups, is highly supportive of an intercontinental transmission event between Asia and Australia, resulting in the local establishment of Clade 2. Furthermore, the only four Australian isolates that did not group with either large clade were located in a small, distinct clade close to Clade 2 and the Asian isolates. All four isolates were recovered in the Australian state of Victoria in 2000 and 2001, suggesting that more than one intercontinental transmission event may have occurred between the neighboring continents of Australia and Asia.

Genomic characterization of Australian ST239 MRSA. To further investigate the molecular differences between the two Australian ST239 clades, as well as understand the impact of the introduction of Clade 2, we estimated the rate of nucleotide substitution, modeled the population size over time, and examined the conservation of protein coding sequences (CDS) among the clades.

Nucleotide substitution rate. An initial global ST239 phylogeny, inferred by maximum-likelihood (ML) phylogenetic analysis, demonstrated a temporal gradient when rooted with a non-ST239 outgroup. This was further supported by the strong correlation between the branch length and the year of isolation ($R^2 = 0.8261$, Fig. 1B). A comparison of the MCC and ML topologies is provided in Fig. S1 in the supplemental material. The Bayesian analysis estimated the mean rate of nucleotide substitutions for the ST239 clone at 1.6×10^{-6} substitutions per site per year ($\text{ss}^{-1} \text{ year}^{-1}$), with a 95% highest posterior density (HPD) interval of 1.2×10^{-6} to $2.0 \times 10^{-6} \text{ ss}^{-1} \text{ year}^{-1}$. This rate is marginally slower than previous estimates of 3.3×10^{-6} (95% confidence interval, 2.5×10^{-6} to 4.0×10^{-6}) $\text{ss}^{-1} \text{ year}^{-1}$ (10), and 3.4×10^{-6} (95% HPD interval, 2.7×10^{-6} to 4.1×10^{-6}) $\text{ss}^{-1} \text{ year}^{-1}$ (12). Our estimate places the emergence of the ST239

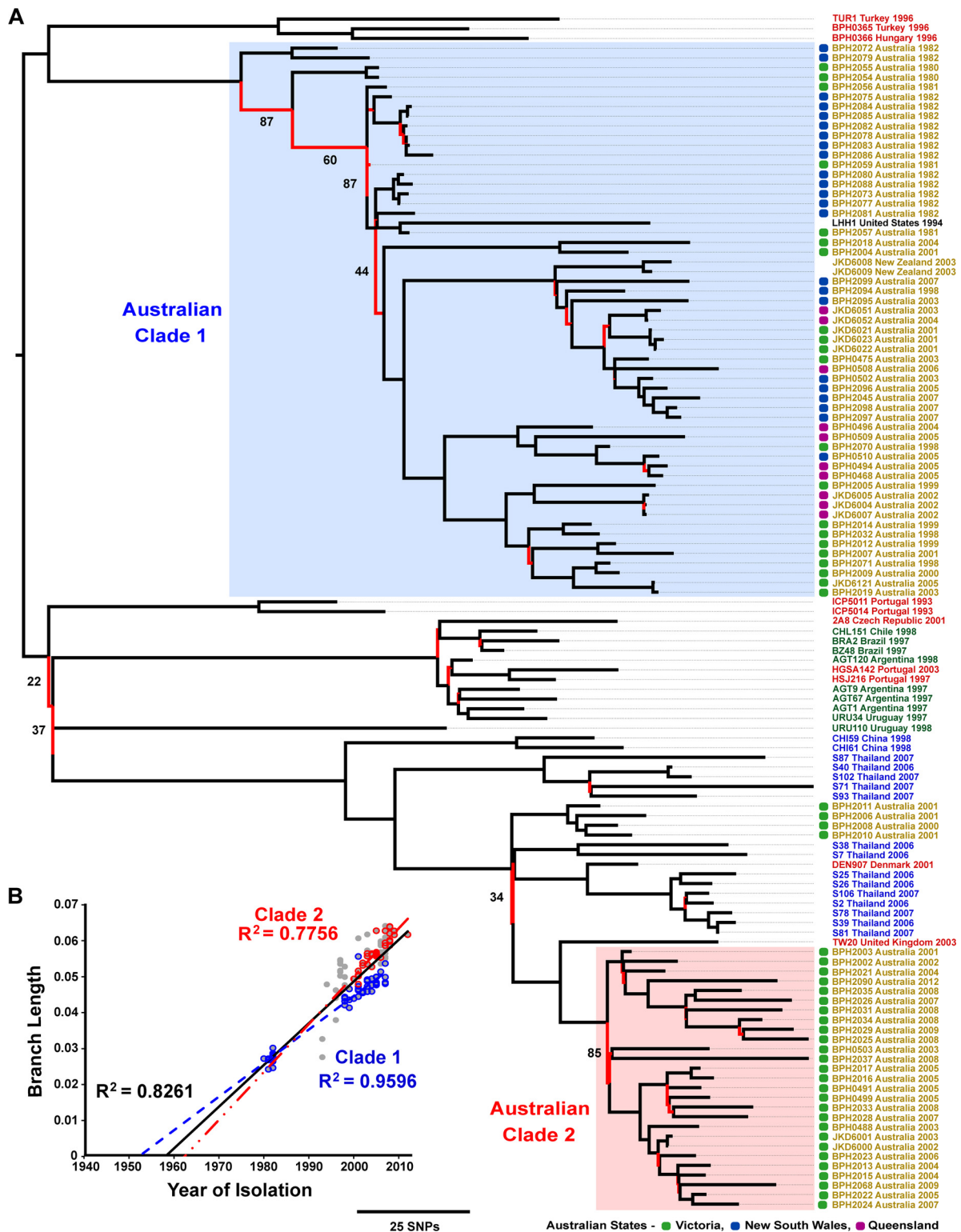


FIG 1 Global phylogeny and molecular clock of ST239 MRSA. (A) MCC tree, generated in BEAST v1.7.5 from 2,716 SNPs in the core genome of 123 isolates, illustrating the global phylogeny of ST239 MRSA. Branches are colored to reflect statistical support; those possessing a posterior probability of $<95\%$ are red. Major branches with $<95\%$ support are annotated with their respective posterior probabilities. Branch annotations provide isolate names, countries, and years of recovery and are colored to reflect the source continents (yellow, Australasia; red, Europe; green, South America; blue, Asia). For Australian isolates, the state of origin is indicated by a colored box (see key). Australian clades are shaded and annotated as Clade 1 and Clade 2. (B) Root-to-tip divergence graph illustrating the strong correlation between year of recovery and branch length, estimated from an ML phylogenetic model. Data points representing Clade 1 and 2 isolates are blue and red, respectively. Trend lines are included for all of the data points (black), Clade 1 (blue), and Clade 2 (red) and have been extrapolated to estimate emergence dates.

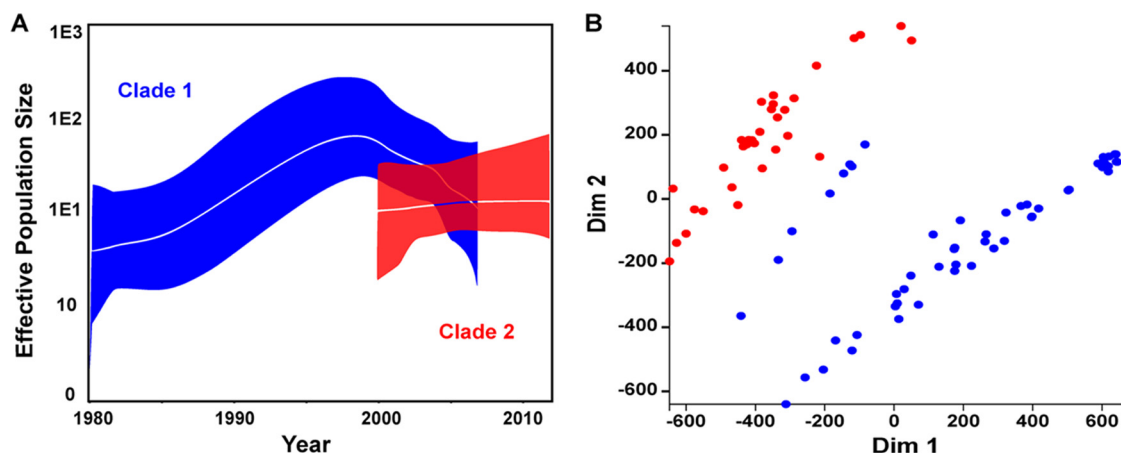


FIG 2 Genomic characterization of Australian ST239 clades. (A) Bayesian Skyline models illustrating the change in the EPS of Clades 1 (blue) and 2 (red). The clades were modeled separately, and their EPSs were overlaid onto a single graph. The median EPS is indicated by the central white line, and the 95% HPD interval is indicated by the colored zone. (B) MDS plot, based on protein ortholog clusters, illustrating the variation in gene content among the Australian isolates from Clades 1 (blue) and 2 (red). Dim, dimension.

clone in 1946 (95% HPD interval, 1885 to 1973). To ascertain the mutation rate and emergence date of each Australian clade, the Bayesian analysis was repeated by testing Clades 1 and 2 independently. Clade 1 was found to mutate at a mean rate of 9.2×10^{-7} (95% HPD interval, 5.8×10^{-7} to 1.3×10^{-6}) $\text{ss}^{-1} \text{year}^{-1}$, and Clade 2 was found to mutate at a mean rate of 1.3×10^{-6} (95% HPD interval, 6.5×10^{-7} to 2.0×10^{-6}) $\text{ss}^{-1} \text{year}^{-1}$. From this, we estimated that Clade 1 emerged in 1962 (95% HPD interval, 1919 to 1981) and Clade 2 emerged in 1986 (95% HPD interval, 1954 to 1998).

Population size. To determine whether Clade 2 had contributed to the persistence of ST239 in the Australian health care environment, a Bayesian Skyline model was employed to estimate changes in the effective population size (EPS) of each clade over time (18). The EPS is estimated from the observed nucleotide variation (genetic diversity) in relation to the mutation rate and can be used to infer changes in the size of a population. These models, illustrated in Fig. 2A, suggested a steady increase in the EPS of Clade 1 from the early 1980s to the late 1990s, followed by a nearly 10-fold decrease over the next 10 years. This result was highly consistent with the reported dominance of the ST239 clone among HA-MRSA isolates in Australia over the last 30 years, as well as the subsequent decrease since 2005 (16). Intriguingly, the EPS of Clade 2 appeared stable throughout the modeled time frame of 2000 to 2012 and could be maintaining the ST239 population in Australia. This may partly explain why the ST239 clone still represents nearly half of HA-MRSA strains encountered in the health care environment nationally and 70% of those in the state of Victoria, where Clade 2 appears to be focused (19).

Gene content variation. To investigate larger genomic differences between the two clades, protein ortholog clustering was used to examine the variation in gene presence or absence. A multidimensional scaling (MDS) representation of these data illustrated discrete, clade-specific clustering (Fig. 2B), further highlighting the genomic divergence in Australian ST239 MRSA. A total of 67 orthologs were identified that differentiated the two clades, each being detected in a minimum of 50% of one clade and absent from all of the isolates of the other clade. It was noted that eight clade 1 isolates, all recovered in the early 1980s, appeared to

share more protein orthologs with Clade 2, as indicated by their location in the MDS plot, than any other Clade 1 isolates. When these eight isolates are excluded, the number of orthologs that differentiate the two clades increased to 209. Of these, 123 (59%) were associated with the three prophages identified in TW20 (ϕSa1 , ϕSa3 , and $\phi\text{Sa}\beta$ -like). Orthologs associated with these prophages were only identified in Clade 2 isolates and the eight Clade 1 isolates. The $\phi\text{Sa}\beta$ -like prophage is of interest in the evolution of the ST239 clone because it carries the recently identified virulence determinant *sasX*. Located at the 3' end of the $\phi\text{SP}\beta$ -like prophage, *sasX* encodes a small, surface-anchored protein that enhances nasal colonization, abscess formation, and tissue damage in the lungs during respiratory infection and has been implicated in the rapid spread of ST239 throughout Asia (13). Li et al. reported that the prevalence of *sasX* among ST239 MRSA in Asia increased substantially between 2003 and 2011, suggesting the preferential uptake of this mobile element. Local alignment identified *sasX* in all of the Clade 2 isolates, as well as the eight clade 1 isolates described above. While this may suggest a role for *sasX* in the evolution of Clade 2, its absence from most Clade 1 isolates indicates that it has not been transferred between the two clades. Furthermore, its presence among only a small number of early Clade 1 isolates suggests that it has been selectively lost, raising questions about the role of *sasX* in the evolution of ST239 MRSA in Australia. Nonetheless, the conservation of *sasX* among Clade 2 isolates is further evidence that this clade originated in Asia.

Collectively, these genomic analyses reveal that while both clades exist and have evolved within the same environment, at a molecular level, they are distinct. These differences are likely to have contributed to the successful expansion of Clade 2 after its reintroduction and the continued persistence of the ST239 clone in this region. These differences also explain some of the phenotypic variation detected in Australian ST239 MRSA.

Changing phenotype of Australian ST239 MRSA. We took advantage of the 32-year time span of our Australian ST239 MRSA culture collection to identify changes in clinically relevant phenotypes (see Fig. S2 in the supplemental material).

Enhanced antimicrobial resistance. Reduced glycopeptide susceptibility has been commonly reported in Australian MRSA

TABLE 1 Comparison of the antimicrobial susceptibility and bacterial fitness of temporal and phylogenetic groups^a

Group (no. of isolates tested)	MIC in $\mu\text{g}/\text{mL}$			Bacterial Fitness
	Vancomycin	Teicoplanin	Daptomycin	Doubling Time (mins)
Isolates recovered during 1980s (19)	1.09	0.81	0.24	39.35
Isolates recovered after 1990 (44)	1.45	1.49	0.51	45.62
	$P = 0.003$	$P < 0.001$	$P < 0.001$	$P < 0.001$
Phylogenetic Clade 1 (38)	1.18	0.97	0.35	43.90
Phylogenetic Clade 2 (25)	1.59	1.79	0.52	43.46
	$P < 0.001$	$P < 0.001$	$P = 0.021$	$P = 0.790$
Clade 1 isolates recovered during 1980s (19)	1.09	0.81	0.24	39.35
Clade 1 isolates recovered after 1990 (19)	1.28	1.17	0.50	48.46
	$P = 0.121$	$P = 0.057$	$P = 0.013$	$P < 0.001$

^a Each value is the mean result of the respective group. *P* values were determined with a Welch two-sample *t* test with log transformations applied to MIC data prior to comparison.

isolates (15, 20). Intermediate-level resistance to vancomycin (vancomycin-intermediate *S. aureus* [VISA]) and teicoplanin was detected at low levels in our collection (15 and 3%, respectively). However, a further 19 isolates (21%) were identified as heterogeneous VISA (hVISA), indicating that more than a third (36%) of the collection had reduced glycopeptide susceptibility. Furthermore, daptomycin nonsusceptibility was detected in 15 isolates (17%). These proportions, although consistent with previous estimations of hVISA and VISA in our region (15), may be an overrepresentation due to the inclusion of a subset of isolates ($n = 24$) with known phenotypes of reduced susceptibility or nonsusceptibility to these agents. They have therefore been excluded from the following analyses of antimicrobial susceptibility and bacterial fitness.

A comparison of antimicrobial susceptibility over time and between phylogenetic clades is presented in Table 1. These analyses demonstrated that the average MIC of vancomycin, teicoplanin, and daptomycin for isolates recovered after 1990 was significantly higher than that for those isolates recovered during the 1980s ($P = 0.003$ to <0.001). A comparison of the average MICs for Clades 1 and 2 also showed a statistically significant difference, with Clade 2 consistently having a higher average MIC of all three antibiotics ($P = 0.016$ to <0.001). As Clade 2 contained only isolates recovered after 2000 and could therefore be creating a bias in the comparison of older and newer isolates, the comparison of temporal groups was repeated by testing Clade 1 independently of Clade 2. While there appeared to be a trend of increasing MICs of all three antibiotics (see Fig. S3 in the supplemental material), consistent with the shift observed in the wider data set, a statistically significant increase was only detected in the average MIC of daptomycin for Clade 1 isolates recovered after 1990 ($P = 0.013$), the *P* values for vancomycin and teicoplanin MICs being low but nonsignificant ($P = 0.121$ and 0.057 , respectively).

There has been significant debate in the literature regarding “MIC creep” in *S. aureus* (21). Previous studies, while demonstrating increasing MICs for MRSA isolates over time, have not performed detailed genotyping to determine if the isolates being tested are related or not. Here, we provide an example of both phenomena contributing to “MIC creep” in one MRSA clone in a single region, adaptive evolution within a clade and the introduction of a new clade, resulting in high rates of VISA infections associated with enhanced morbidity and poor treatment responses (20, 22, 23).

GWAS. To identify the genetic mediators of enhanced antimicrobial resistance, in particular to vancomycin, a genome-wide association study (GWAS) approach was implemented. These types of studies often require large numbers of isolates to detect associations, as the significance threshold is adjusted to account for multiple testing (critical *P* value, $\leq 0.05/n$). Therefore, it was unsurprising that an initial comparison of the frequency of all of the core genome SNPs resulting in nonsynonymous amino acid changes between isolates with vancomycin MICs of ≤ 2 and $\geq 3 \mu\text{g}/\text{mL}$ (Fig. 3A) failed to detect any significant associations. Interestingly, the smallest *P* value achieved in this analysis was that of an SNP resulting in an RpoB_{H481Y} amino acid substitution ($P = 0.00029$). This exact mutation within the beta subunit of the RNA

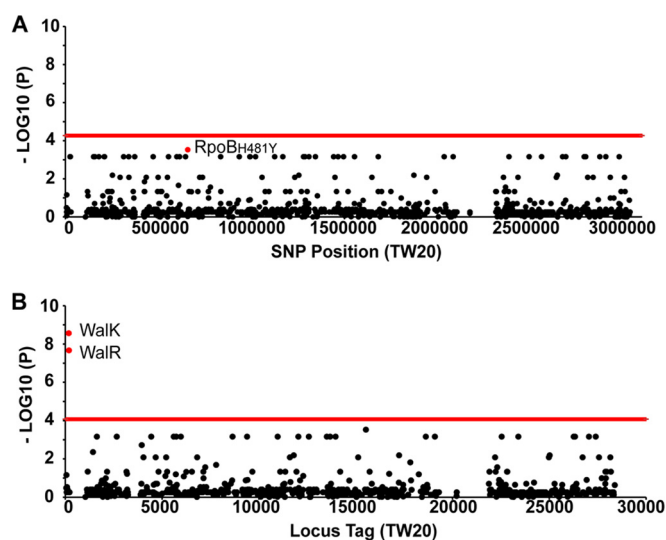


FIG 3 GWAS of reduced vancomycin susceptibility in ST239 MRSA. Manhattan plots illustrate the results of two GWASs comparing isolates with vancomycin MICs of ≤ 2 or $\geq 3 \mu\text{g}/\text{mL}$. The red line indicates the threshold of significance (*P* value of $\leq 0.05/\text{number of hypotheses tested}$), as determined by chi-square analysis. (A) Manhattan plot illustrating the associations between SNPs and vancomycin MICs. The single red data point represents the SNP resulting in an RpoB_{H481Y} amino acid substitution. (B) Second Manhattan plot illustrating the associations of CDSs containing a SNP(s), independent of the SNP location within the locus, with vancomycin MICs. Highlighted red data points represent the *walkK* and *walR* genes, the only significant associations identified in this analysis.

polymerase has previously been demonstrated to lead to reduced vancomycin susceptibility by using genetic manipulation experiments (24) and was also identified in a recent GWAS of reduced glycopeptide susceptibility (25). In the Australian ST239 collection, this SNP was identified in three phylogenetically dispersed isolates with high vancomycin MICs. Furthermore, this analysis revealed a high number of SNPs within the essential two-component regulator *walKR*, all with very low *P* values ($P = 0.00069$).

To identify whether *walKR*, or any other gene, is a hot spot for mutations associated with increased vancomycin MICs, a modified GWAS was performed. In this analysis, each CDS was classified as either “unaffected” or “affected”, the latter being assigned if a SNP was identified within the locus, regardless of its location. The frequencies of “affected” CDS in the same MIC groups described above were then compared (Fig. 3B). This showed that variant sequences of both *walK* and *walR* were highly associated with increased vancomycin MICs ($P = 2.721 \times 10^{-9}$ and 2.149×10^{-8} , respectively). However, the exact mutations were highly variable, with three different *walR* mutations identified in four isolates and seven different *walK* mutations identified in seven isolates (see Table S1 in the supplemental material). The DNA sequences of *rpoB* and *walKR* are highly conserved in *S. aureus*; however, mutations in both loci have previously been linked to not only enhanced glycopeptide resistance (26) but also alterations in virulence and, in the case of *rpoB*, resistance to the host innate immune responses (27). The selective pressures driving resistance to last-line antimicrobials in Australian ST239 MRSA are therefore also likely to be having an impact on the fitness and virulence of this clone.

Relative bacterial fitness. To determine whether there has also been a change in the relative fitness of Australian ST239 MRSA, the average replication rate of each isolate was measured (see Fig. S2 in the supplemental material). The mean doubling time of the Australian collection ($n = 87$) was 46.51 ± 8.48 min or 43.73 ± 6.61 min when the subset of isolates ($n = 24$) with known resistance phenotypes was excluded. This finding is similar to the previously published replication rate of five Australian ST239 isolates of 42.2 ± 6.5 min (28). A comparison of older and newer isolates identified a statistically significant increase in the mean doubling time of isolates recovered after 1990 ($P < 0.001$) (Table 1). This increase was also observed among newer Clade 1 isolates when they were tested independently ($P < 0.001$). No significant difference between the mean doubling times of Clades 1 and 2 was observed. However, Clade 2 isolates did demonstrate a significantly slower replication rate than Clade 1 isolates recovered in the 1980s ($P = 0.017$) but a significantly faster rate than those recovered after 1990 ($P = 0.008$). This suggested that reduced fitness among Australian ST239 MRSA was largely due to an increase in the replication rate of Clade 1. Furthermore, Clade 2, while demonstrating enhanced antimicrobial resistance, has not suffered the same loss of fitness, another possible contributor to its successful establishment within the Australian health care environment.

Evolution to attenuated virulence but preserved *in vivo* persistence in ST239 MRSA. Reduced virulence has been demonstrated in HA-MRSA clones, particularly those that are multidrug resistant (29–31); however, the evolution of virulence over time has not been investigated. To determine the impact of hospital adaptation on the virulence of Australian ST239 MRSA, the activ-

ity of the accessory gene regulator (*agr*), the primary global virulence regulatory system in *S. aureus*, was measured, and the results were confirmed in a murine infection model.

δ -Hemolysin production. The activity of the *agr* complex was assessed by determining the phenotypic production of δ -hemolysin (32). Of the 87 isolates in the collection, 40 (46%) demonstrated δ -hemolysis and were thus considered to have an active *agr* locus. The remaining 47 (54%) did not demonstrate δ -hemolysis and were therefore considered *agr* defective. A comparison of *agr* activity and the year of isolation showed a dramatic shift in the proportion of isolates with an active *agr* system over time. Nearly 84% (16/19) of the isolates recovered during the 1980s appeared to have an active *agr* locus. However, this dropped to 52% (12/23) of the isolates recovered between 1998 and 2002 and 31% (11/36) of those recovered between 2003 and 2007. All of the isolates recovered after 2007 appeared to be *agr* defective, although this time frame has been sparsely sampled. This suggests strong selection pressure against *agr* activity in the Australian health care environment and is consistent with the recent work of Paulander et al. (33) in which exposure to selected antimicrobials resulted in increased *agr* expression and reduced bacterial fitness. They proposed that certain antibiotics may be indirectly driving the selection of *agr*-defective isolates by imposing an RNAPIII-associated fitness cost on isolates with functional loci. When the *agr* activities of the phylogenetic clades were compared, it became apparent that all Clade 2 isolates were potentially *agr* defective. When Clade 2 is excluded, a decrease in *agr* activity over time is still observed, the percentage of Clade 1 isolates with an active *agr* locus dropping from 84% to 75% (12/16) and then to 52% (11/21) for the pre-1990, 1998-to-2002, and 2003-to-2007 time frames, respectively.

Reduction of *agr* activity is most commonly associated with mutations in the *agrABCD* locus (32, 34). In the ST239 Australian collection, only eight SNPs and two single base insertions were identified within *agrABCD* but only seven of these resulted in nonsynonymous amino acid changes (see Table S2 in the supplemental material). Intriguingly, a total of 37 isolates were identified with a defective *agr* system but there was no mutation in *agrABCD* to explain the loss of function. When the presence of *agrABCD* mutation was compared to our regional phylogeny, it became apparent that the most common SNPs, resulting in predicted Agr_{R6S} and Agr_{R393C} changes, were found only within Clade 1 isolates (see Fig. S2 in the supplemental material). The predicted Agr_{R6S} change was identified in all Clade 1 isolates, regardless of *agr* activity. Alternatively, the Agr_{R393C} change was identified in a subset of five isolates, all *agr* defective and all belonging to a single subclade of Clade 1. The majority of the isolates that we observed with a defective *agr* system and no *agr* mutations identified belonged to Clade 2.

RNAPIII expression. To confirm these findings, reverse transcription-quantitative PCR was performed with a subset of isolates randomly selected from three groups, Clade 1 isolates recovered during the 1980s, Clade 1 isolates recovered after 1990, and Clade 2 isolates. RNAPIII expression among the representative Clade 1 isolates revealed a dramatic reduction in the expression of the *agr* complex over time (Fig. 4A), consistent with the findings of the δ -hemolysin assay. Isolates recovered in the 1980s showed very high RNAPIII expression, often above that of the strong positive control JKD6159, a highly virulent ST93 community MRSA isolate from Australia that produces large amounts of

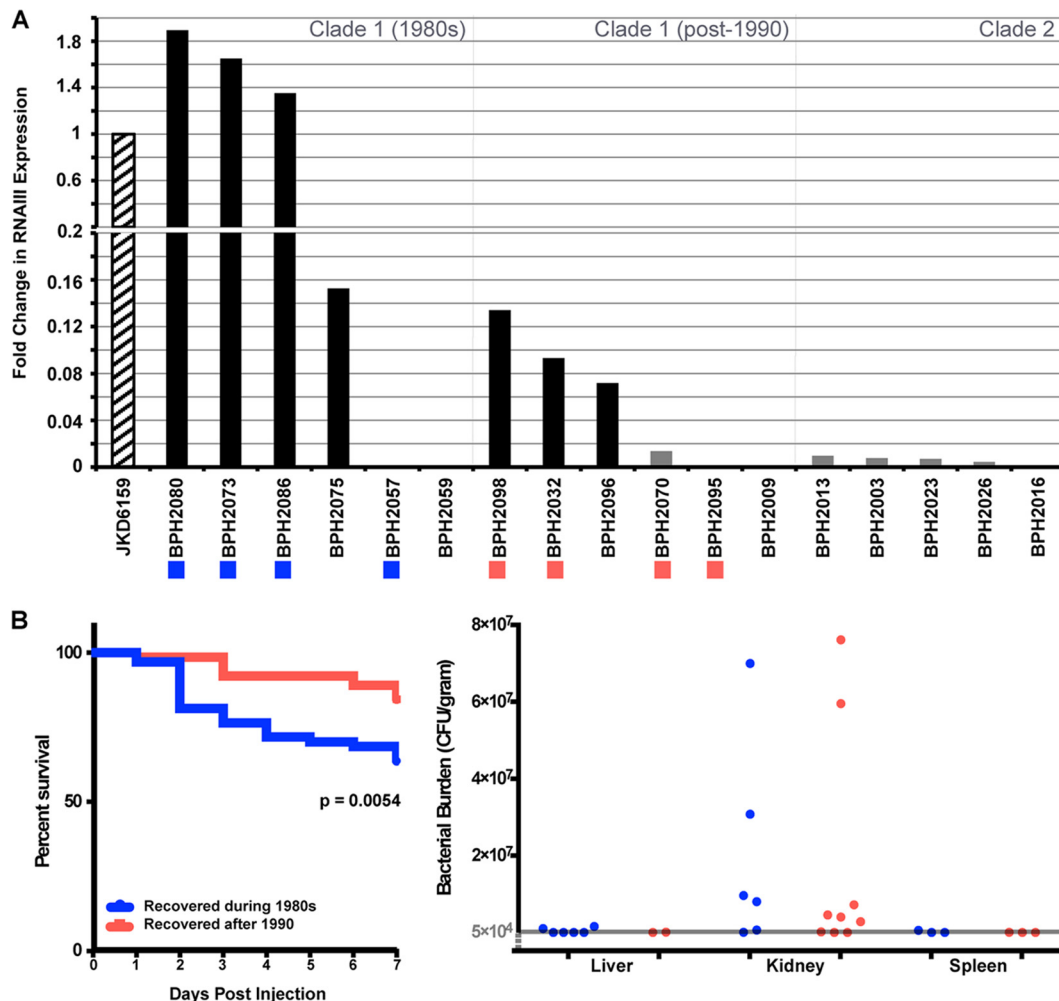


FIG 4 Evolution of ST239 MRSA toward attenuated virulence. (A) Fold changes in RNAIII expression levels compared to control JKD6159, as determined by reverse transcription-quantitative PCR. Bar color indicates activity of the AGR complex as determined by the delta-hemolysin assay (black, active; grey, inactive). Colored boxes below isolate names indicate those isolates that were tested in the *in vivo* murine sepsis model; colors represent the temporal groups to which they belong (see panel B). (B) Results obtained with the *in vivo* sepsis model. (Left) Seven-day survival curves grouped by year of isolation as indicated in the key. Each group represents four bacterial isolates, each tested in 16 animals. The *P* value is for a comparison of Kaplan-Meier curves with a log-rank test. (Right) Bacterial burdens in the kidneys at 7 days postinjection. Each data point represents one animal infected with one isolate. The detection threshold is shown by the grey line; only values above this level are displayed.

α -hemolysin (35), a product positively regulated by RNAIII. The discovery of isolates with expression levels higher than that of JKD6159 suggested that early ST239 isolates were potentially highly virulent but extended evolution within the hospital environment has led to less virulent phenotypes. Of these 12 isolates, only one produced a result inconsistent with the δ -hemolysin assay. BPH2070 showed no δ -hemolysis but expressed RNAIII at low levels. This suggests that a negative result in the initial δ -hemolysin screen may be indicative of both negative and very-low-level RNAIII expression. Of the four Clade 1 isolates that did not express RNAIII, only BPH2095 had an *agr* mutation that could account for the loss of activity, carrying a single base insertion within *agrC* resulting in a reading frameshift. No Clade 2 isolates demonstrated δ -hemolysis; however, four of the five representative isolates expressed RNAIII at low levels, much like BPH2070. The single Clade 2 isolate (BPH2016) that did not express RNAIII carried a single point mutation with a predicted

*AgrC*_{P24S} amino acid change. These results suggest that most Clade 2 isolates that do not possess an *agrABCD* mutation may express RNAIII at low levels. Furthermore, it highlights that at least three different mechanisms are contributing to attenuated virulence in the ST239 clone, mutation within the locus disrupting its activity, as well as external factors leading to both complete inactivity and reduced RNAIII expression.

***In vivo* sepsis model.** To determine whether Australian ST239 MRSA was indeed evolving toward attenuated virulence, eight of the 12 Clade 1 isolates, four from each temporal group, tested for RNAIII expression were also assessed in an *in vivo* murine sepsis model. As the reduction in RNAIII expression over time was observed only in Clade 1, Clade 2 isolates were not tested in this model. The results of the *in vivo* murine sepsis model demonstrated that Clade 1 isolates recovered in the 1980s were significantly more virulent (Fig. 4B), in terms of both survival ($P = 0.0054$) and weight loss at day 2 ($P < 0.001$), day 3 ($P = 0.046$),

and day 4 ($P = 0.0056$), compared to newer clade 1 isolates (see Fig. S4 in the supplemental material), confirmation of the evolution of this clade toward attenuated virulence. Additionally, it has been well demonstrated that RNAIII expression contributes to virulence (35). Consistent with this, when isolates were regrouped on the basis of RNAIII expression, those with high-level expression were more virulent in the murine model than isolates with moderate or no RNAIII expression, as shown by both survival ($P = 0.0027$) and weight loss at day 2 ($P < 0.0001$) and day 4 ($P = 0.0434$) (see Fig. S4 in the supplemental material). Interestingly, there appeared to be no difference in virulence between isolates with moderate RNAIII expression and those with no RNAIII expression. The regulation of virulence in *S. aureus* is multifactorial, and these results could indicate that lower-level RNAIII expression may be insufficient in these isolates for the transcriptional regulation of downstream virulence determinants. Alternatively, isolates with defective *agr* loci may have compensatory mechanisms of pathogenesis that are independent of *agr* and have not been explored in this study.

Previous studies have shown that isolates with attenuated virulence are capable of persistent infection (27), with a viable bacterial cell count detectable in the kidneys at 7 days postinjection in a murine sepsis model. We also noted that our isolates were capable of persistence; however, it was not associated with a specific phenotypic group (Fig. 4B). Our results suggest that ST239 MRSA is capable of persistent infection regardless of its relative virulence level or evolutionary age.

Conclusions. Using detailed genomic and phenotypic analyses, this study has revealed a very clear picture of the localized evolution of one of the most important global clones of HA-MRSA. The high resolving power of bacterial genomics has uncovered a previously unrecognized reimportation of ST239 MRSA into Australia, perpetuating the ST239 epidemic in our region, and demonstrated the genomic diversity between the cocirculating clades. Changes common to these two independent clades have revealed an adaptation toward increased antimicrobial resistance and a HA-MRSA clone that is primed for higher-level glycopeptide and daptomycin resistance. The apparent cost of these changes has been a significant reduction in the replication rate and virulence measured by analysis of doubling times, *agr* activity, and a murine sepsis model. However, these changes may actually represent coselected adaptive advantages, as the capacity of the ST239 clone to cause persistent invasive infection remains significant. These results shed new light on the complex evolution of *S. aureus* clones that predominate within the health care environment and illustrate how projects utilizing both comparative genomic and functional approaches are essential to understanding the impact of current antibiotic usage policies on hospital pathogens.

MATERIALS AND METHODS

Bacterial isolates. This study utilized a temporal and geographically dispersed collection of 123 ST239 *S. aureus* isolates. The culture collection consisted of 87 isolates recovered from the Australia-New Zealand region between 1980 and 2012. For genomic analyses, this collection was supplemented with the genome sequences of an additional 35 isolates from the international collection of Harris et al. (10). Relevant isolate demographics are available in Data Set S1 in the supplemental material.

Genome sequencing and analyses. Of the Australian isolates, 71 were sequenced in this study with MiSeq (Illumina) and a read length of 2×150 bp. The genome sequence reads of previously published strains are available under the following accession numbers: TW20, GenBank acces-

sion no. FN433596; the 35 international ST239 strains, Sequence Read Archive accession no. ERA000102; JKD6008, GenBank accession no. CP002120. Genomic and phylogenetic analyses were performed as described in Text S1 in the supplemental material.

Phenotypic analyses. The experimental methodology and analysis of antimicrobial susceptibility, growth rate, functionality of the *agr* locus and the *in vivo* murine sepsis model are described in Text S1 in the supplemental material. All animal studies were performed in accordance with the Animal Research Ethics Committee at the University of Melbourne.

Nucleotide sequence accession number. The sequence reads determined in this study were submitted to the European Nucleotide Archive (ENA) under study number ERP009308, and can be accessed via <http://www.ebi.ac.uk/ena/data/view/ERP009308>.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00080-15/-/DCSupplemental>.

Text S1, PDF file, 0.3 MB.

Data Set S1, PDF file, 0.1 MB.

Figure S1, PDF file, 0.03 MB.

Figure S2, PDF file, 0.1 MB.

Figure S3, PDF file, 0.1 MB.

Figure S4, PDF file, 0.1 MB.

Table S1, PDF file, 0.1 MB.

Table S2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

This work was funded by the National Health and Medical Research Council, Australia (fellowships APP1023526 to B.P.H., APP1008549 to T.P.S., and APP1061409 to K.E.H.).

We thank Elizabeth Grabsch of the Microbiology Department of Austin Health for assistance with the initial pulsed-field gel electrophoresis characterization of isolates.

REFERENCES

- Wyllie D, Paul J, Crook D. 2011. Waves of trouble: MRSA strain dynamics and assessment of the impact of infection control. *J Antimicrob Chemother* 66:2685–2688. <http://dx.doi.org/10.1093/jac/ckr392>.
- Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc Natl Acad Sci U S A* 99:7687–7692. <http://dx.doi.org/10.1073/pnas.122108599>.
- McAdam PR, Templeton KE, Edwards GF, Holden MT, Feil EJ, Aanensen DM, Bargawi HJ, Spratt BG, Bentley SD, Parkhill J, Enright MC, Holmes A, Girvan EK, Godfrey PA, Feldgarden M, Kearns AM, Rambaut A, Robinson DA, Fitzgerald JR. 2012. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 109:9107–9112. <http://dx.doi.org/10.1073/pnas.1202869109>.
- Robinson DA, Enright MC. 2004. Evolution of *Staphylococcus aureus* by large chromosomal replacements. *J Bacteriol* 186:1060–1064. <http://dx.doi.org/10.1128/JB.186.4.1060-1064.2004>.
- Pavillard R, Harvey K, Douglas D, Hewstone A, Andrew J, Collopy B, Asche V, Carson P, Davidson A, Gilbert G, Spicer J, Tosolini F. 1982. Epidemic of hospital-acquired infection due to methicillin-resistant *Staphylococcus aureus* in major Victorian hospitals. *Med J Aust* 29:451–454.
- Cookson BD, Phillips I. 1988. Epidemic methicillin-resistant *Staphylococcus aureus*. *J Antimicrob Chemother* 21:57–65. http://dx.doi.org/10.1093/jac/21.suppl_C.57.
- Dubin DT, Chikramane SG, Inglis B, Matthews PR, Stewart PR. 1992. Physical mapping of the *mec* region of an Australian methicillin-resistant *Staphylococcus aureus* lineage and a closely related American strain. *J Gen Microbiol* 138:169–180. <http://dx.doi.org/10.1099/00221287-138-1-169>.
- Rountree PM, Beard MA. 1968. Hospital strain of *Staphylococcus aureus* with particular reference to methicillin-resistant strains. *Med J Aust* 2:1163–1168.
- Smyth DS, McDougal LK, Gran FW, Manoharan A, Enright MC, Song J-H, de Lencastre H, Robinson DA. 2010. Population structure of a

- hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. PLoS One 5:e8582. <http://dx.doi.org/10.1371/journal.pone.0008582>.
10. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327: 469–474. <http://dx.doi.org/10.1126/science.1182395>.
 11. Castillo-Ramírez S, Corander J, Marttinen P, Aldeljawy M, Hanage WP, Westh H, Boye K, Bentley SD, Parkhill J, Holden MT, Feil EJ, Feil EJ. 2012. Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol* 13: R126. <http://dx.doi.org/10.1186/gb-2012-13-12-r126>.
 12. Gray RR, Tatem AJ, Johnson JA, Alekseyenko AV, Pybus OG, Suchard MA, Salemi M. 2011. Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a Bayesian framework. *Mol Biol Evol* 28:1593–1603. <http://dx.doi.org/10.1093/molbev/msq319>.
 13. Li M, Du X, Villaruz AE, Diep BA, Wang D, Song Y, Tian Y, Hu J, Yu F, Lu Y, Otto M. 2012. MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med* 18:816–819. <http://dx.doi.org/10.1038/nm.2692>.
 14. Espedido BA, Steen JA, Barbagiannakos T, Mercer J, Paterson DL, Grimmond SM, Cooper MA, Gosbell IB, van Hal SJ, Jensen SO. 2012. Carriage of an ACME II variant may have contributed to methicillin-resistant *Staphylococcus aureus* sequence type 239-like strain replacement in Liverpool Hospital, Sydney, Australia. *Antimicrob Agents Chemother* 56:3380–3383. <http://dx.doi.org/10.1128/AAC.00013-12>.
 15. Horne KC, Howden BP, Grabsch EA, Graham M, Ward PB, Xie S, Mayall BC, Johnson PD, Grayson ML. 2009. Prospective comparison of the clinical impacts of heterogeneous vancomycin-intermediate methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-susceptible MRSA. *Antimicrob Agents Chemother* 53:3447–3452. <http://dx.doi.org/10.1128/AAC.01365-08>.
 16. Australian Group on Antibiotic Resistance. 2004, posting date. *Staphylococcus aureus* programme (SAP) hospital/community surveys, MRSA epidemiology and typing reports: 2003–2011. Australian Group on Antimicrobial Resistance, Perth, WA, Australia. <http://www.agargroup.org/surveys>.
 17. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MVN, Anderson TL, Roberts SA, Warren SJC, Coombs GW, Tan H.L., Gao W, Johnson PDR, Howden BP. 2014. Genetic and molecular predictors of high vancomycin minimum inhibitory concentration in *Staphylococcus aureus* bacteremia isolates. *J Clin Microbiol* 52:3384–3393. <http://dx.doi.org/10.1128/JCM.01320-14>.
 18. Minin VN, Bloomquist EW, Suchard MA. 2008. Smooth Skyride through a rough Skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* 25:1459–1471. <http://dx.doi.org/10.1093/molbev/msn090>.
 19. Coombs GW, Nimmo GR, Pearson JC, Collignon PJ, Bell JM, McLaws ML, Christiansen KJ, Turnidge JD, Australian Group on Antimicrobial Resistance. 2013. Australian Group on Antimicrobial Resistance hospital-onset *Staphylococcus aureus* surveillance programme annual report, 2011. *Commun Dis Intell Q Rep* 37:E210–E218.
 20. Howden BP, Davies JK, Johnson PD, Stinear TP, Grayson ML. 2010. Reduced vancomycin susceptibility in *Staphylococcus aureus*, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: resistance mechanisms, laboratory detection, and clinical implications. *Clin Microbiol Rev* 23:99–139. <http://dx.doi.org/10.1128/CMR.00042-09>.
 21. van Hal SJ, Fowler VG, Jr. 2013. Is it time to replace vancomycin in the treatment of MRSA infections? *Clin Infect Dis* 56:1779–1788. <http://dx.doi.org/10.1093/cid/cit178>.
 22. Howden BP, Ward PB, Charles PG, Korman TM, Fuller A, du Cros P, Grabsch EA, Roberts SA, Robson J, Read K, Bak N, Hurley J, Johnson PD, Morris AJ, Mayall BC, Grayson ML. 2004. Treatment outcomes for serious infections caused by methicillin-resistant *Staphylococcus aureus* with reduced vancomycin susceptibility. *Clin Infect Dis* 38:521–528. <http://dx.doi.org/10.1086/381202>.
 23. Charles PG, Ward PB, Johnson PD, Howden BP, Grayson ML. 2004. Clinical features associated with bacteremia due to heterogeneous vancomycin-intermediate *Staphylococcus aureus*. *Clin Infect Dis* 38: 448–451. <http://dx.doi.org/10.1086/381093>.
 24. Matsuo M, Hishinuma T, Katayama Y, Cui L, Kapi M, Hiramatsu K. 2011. Mutation of RNA polymerase β subunit (*rpoB*) promotes hVISA-to-VISA phenotypic conversion of strain Mu3. *Antimicrob Agents Chemother* 55:4188–4195. <http://dx.doi.org/10.1128/AAC.00398-11>.
 25. Alam MT, Petit RA III, Crispell EK, Thornton TA, Conneely KN, Jiang Y, Satola SW, Read TD. 2014. Dissecting vancomycin-intermediate resistance in *Staphylococcus aureus* using genome-wide association. *Genome Biol Evol* 6:1174–1185. <http://dx.doi.org/10.1093/gbe/evu092>.
 26. Howden BP, McEvoy CR, Allen DL, Chua K, Gao W, Harrison PF, Bell J, Coombs G, Bennett-Wood V, Porter JL, Robins-Browne R, Davies JK, Seemann T, Stinear TP. 2011. Evolution of multidrug resistance during *Staphylococcus aureus* infection involves mutation of the essential two component regulator WalKR. *PLoS Pathog* 7:e1002359. <http://dx.doi.org/10.1371/journal.ppat.1002359>.
 27. Gao W, Cameron DR, Davies JK, Kostoulas X, Stepnell J, Tuck KL, Yeaman MR, Peleg AY, Stinear TP, Howden BP. 2013. The *rpoB* H₄₈₁Y rifampicin resistance mutation and an active stringent response reduce virulence and increase resistance to innate immune responses in *Staphylococcus aureus*. *J Infect Dis* 207:929–939. <http://dx.doi.org/10.1093/infdis/jis772>.
 28. Okuma H, Iwakawa K, Turnidge JD, Grubb WB, Bell JM, O'Brien FG, Coombs GW, Pearman JW, Tenover FC, Kapi M, Tiensasitorn C, Ito T, Hiramatsu K. 2002. Dissemination of new methicillin-resistant *Staphylococcus aureus* clones in the community. *J Clin Microbiol* 40:4289–4294. <http://dx.doi.org/10.1128/JCM.40.11.4289-4294.2002>.
 29. Rudkin JK, Edwards AM, Bowden MG, Brown EL, Pozzi C, Waters EM, Chan WC, Williams P, O'Gara JP, Massey RC. 2012. Methicillin resistance reduces the virulence of healthcare-associated methicillin-resistant *Staphylococcus aureus* by interfering with the *agr* quorum sensing system. *J Infect Dis* 205:798–806. <http://dx.doi.org/10.1093/infdis/jir845>.
 30. Howden BP, Johnson PD, Ward PB, Stinear TP, Davies JK. 2006. Isolates with low-level vancomycin resistance associated with persistent methicillin-resistant *Staphylococcus aureus* bacteremia. *Antimicrob Agents Chemother* 50:3039–3047. <http://dx.doi.org/10.1128/AAC.00422-06>.
 31. Sakoulas G. 2006. The accessory gene regulator (*agr*) in methicillin-resistant *Staphylococcus aureus*: role in virulence and reduced susceptibility to glycopeptide antibiotics. *Drug Discov Today Dis Mech* 3:287–294. <http://dx.doi.org/10.1016/j.ddmec.2006.06.007>.
 32. Herbert S, Ziebandt A.-K, Ohlsen K, Schäfer T, Hecker M, Albrecht D, Novick R, Götz F. 2010. Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. *Infect Immun* 78:2877–2889. <http://dx.doi.org/10.1128/IAI.00088-10>.
 33. Paulander W, Nissen Varming A, Bæk KT, Haaber J, Frees D, Ingmer H. 2013. Antibiotic-mediated selection of quorum-sensing-negative *Staphylococcus aureus*. *mBio* 3:e00459–e00412. <http://dx.doi.org/10.1128/mBio.00459-12>.
 34. Moore PC, Lindsay JA. 2001. Genetic variation among hospital isolates of methicillin-sensitive *Staphylococcus aureus*: evidence for horizontal transfer of virulence genes. *J Clin Microbiol* 39:2760–2767. <http://dx.doi.org/10.1128/JCM.39.8.2760-2767.2001>.
 35. Chua KY, Monk IR, Lin Y.-H, Seemann T, Tuck KL, Porter JL, Stepnell J, Coombs GW, Davies JK, Stinear TP, Howden BP. 2014. Hyperexpression of α -hemolysin explains enhanced virulence of sequence type 93 community-associated methicillin-resistant *Staphylococcus aureus*. *BMC Microbiol* 14:31. <http://dx.doi.org/10.1186/1471-2180-14-31>.

2.2.2 Supplementary Materials

Baines, S. L., K. E. Holt, M. B. Schultz, T. Seemann, B. O. Howden, S. O. Jensen, S. J. van Hal, G. W. Coombs, N. Firth, D. R. Powell, T. P. Stinear and B. P. Howden (2015). "Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*." MBio; DOI: 10.1128/mBio.00080-15.

Available in Appendix 1: Dataset S1 Summary of study isolates and genome sequencing. The relevant demographic and whole-genome sequencing information of all isolates included in this study are shown.

Text S1: Supplementary Materials & Methods

Bacterial Strains & Storage.

This study utilised a temporal and geographically dispersed collection of 123 ST239 *S. aureus* isolates. Of these, 87 were recovered from the Australian states of Victoria (n=52), New South Wales (n=10), and Queensland (n=10), as well as New Zealand (n=2), between 1980 and 2012. These isolates were initially identified as ST239 using PFGE, or an ST239 specific PCR as previously described (304). Twenty-seven isolates are from previously published studies (170, 305, 306), and the remaining 60 are novel. The majority of these represent clinical isolates causing either septicaemia (n=50), or tissue infection (n=7) in non-duplicate patients. Twelve represent paired, or triplet, isogenic isolates recovered from patients who developed reduced glycopeptide susceptible variants during persistent infection. The remaining 17 isolates, all recovered in the early 1980s, had no clinical data available. This collection was supplemented with the genome sequences of an additional 35 international ST239 isolates (300). Bacterial strains were stored in glycerol broth at -80°C and sub-cultured twice on horse blood agar (Oxoid) prior to all testing. If liquid culture was required, the second sub-culture was performed in heart infusion broth.

Genome Sequencing & Analysis.

A total of 71 isolates were sequenced in this study. Genomic DNA was extracted using the GenElute Genomic DNA Kit (Sigma), multiplex short-read DNA libraries were created with the Nextera XT DNA Preparation Kit (Illumina), and whole genome sequencing was performed on the MiSeq (Illumina) using 2 x 150 bp chemistry; all performed as per the manufacturer's recommendations. The genome sequence reads of previously published strains are available under the following accession numbers: TW20 (GenBank, FN433596), 35 international ST239 strains (Sequence Read Archive, ERA000102), JKD6008 (GenBank, CP002120). All other sequences were submitted to the National Centre for Biotechnology Information GenBank and are available under BioProject SAMN03083454.

A read mapping approach was used to align short-read sequences of all 123 strains to reference ST239 strain TW20 using SHRiMP v2.0 (307). Identification of mutation in read alignments were performed as previously described (308).

Detection of Recombination. Core genome SNPs potentially introduced by recombination were identified using a modified method of that proposed by Croucher *et. al.* (309). A phylogeny was established using RAxML (310),

then SNPs were mapped onto the branches using PAML v4 (311); each branch representing a unique group of temporally related SNPs. Each group was examined for spatial clustering using a chi-squared analysis with a Bonferroni correction applied to determine significance. Clustered SNPs were removed, and the analysis repeated with no further clustering identified.

Phylogenomic Analyses. An initial global population model was estimated using PhyML v3 (312), under a GTR model of nucleotide substitution with 1000 bootstrap replicates and rooted with the non-ST239 *S. aureus* strain FPR3757 (GenBank, NC_007793.1). This phylogeny was used to assess the correlation between branch length and year of isolation with Path-o-Gen v1.3 (313).

Phylogenetic analysis of the core genome, non-recombinant SNPs was completed using BEAST v1.7.5 (refer to <http://beast.bio.ed.ac.uk> for all software in this paragraph), also using the GTR model. Three molecular clock models were initially trialled (strict clock, uncorrelated relaxed log normal (UCLN) and random local clock) with the best likelihood achieved using the UCLN model (unpublished data), a result consistent with Gray *et al.* (299). A constant size tree prior was used for modelling phylogeny and a Bayesian Skyline model for estimating the effective population size; run with the default parameters of a piecewise-constant model and 10 groups. For the global phylogeny (FIG 1A) three 100 million Monte Carlo Markov Chain (MCMC) analyses were run and combined, and for individual clade analysis four 10 million MCMC chains were run and combined, sampling every 1000th steps to achieve effective sampling. The burnin was set at the earliest point at which a stable likelihood was achieved in each run (interpreted visually in Tracer v1.5). No runs were discarded as the marginal likelihoods of all iterations was ≤ 2 (14). Runs were combined using LogCombiner v1.7.5 and re-sampled at a frequency that would achieve approximately 10,000 samples. TreeAnnotator v1.7.5 was used to select the topology with MCC and annotate the tree with mean heights. Final alterations to tree topology, including branch width, colour and annotation, were undertaken in FigTree v1.4.2. The model of population size was generated in Tracer v 1.5 using the Bayesian Skyline reconstruction tool. Calculation of the molecular clock as SNPs per site per year was completed by multiplying the rate estimated in BEAST with the number of variant sites tested, then dividing the total by the approximate size of a *S. aureus* genome (2.8Mb) (299).

Gene Presence/Absence. Sequence reads were de novo assembled as previously described (308), and the resulting contigs were used to construct an orthologous protein matrix with proteinortho v5 (314), using blastp+ with a default e-value of 1×10^{-5} . This matrix was used to generate a multidimensional scaling plot, in FriPan (<http://drpowell.github.io/FriPan/>, unpublished), representing the variation in gene content between strains (FIG 2B).

Phenotypic Analysis.

Antimicrobial Susceptibility Testing. The minimum inhibitory concentration for each strain were determined for vancomycin, teicoplanin and daptomycin, using Etests (bioMérieux), following the manufacturer's recommendations. Results were interpreted following the Clinical Laboratory Standards Institute (CLSI) guidelines (315). To determine reduced glycopeptide susceptibility Macro-Etests, using a 2.0 McF solution, were performed as previously described by Walsh *et al.* (316). *S. aureus* ATCC25923 and Mu3 were included as controls.

Genome Wide Association Study. Allelic association was performed using the GWAS software Plink (317). Phenotypic groups were defined by a vancomycin MIC of $\leq 2 \mu\text{g/mL}$ or $\geq 3 \mu\text{g/mL}$. Core genome SNPs that resulted in non-synonymous amino acid changes were compared testing each SNP individually as well as grouping SNPs based on location and assigning an allele value for CDS with or without mutation. A Bonferroni correction was used to account for multiple hypothesis testing; the critical p-value determined as $\leq 0.05/\text{number of SNPs or CDSs tested}$.

Growth Curves. The average doubling time for each strain was determined using growth curves. These were performed as previously described (318), with the one exception that Mueller Hinton broth (BD) was used. Strains were tested in both biological and technical triplicate. To convert optical density into doubling time, technical replicates were averaged and graphed on a log scale. The OD600 at the last time point in which all strains entered exponential growth (40 minutes) and first time point at which the first strain entered stationary phase (130 minutes) was used as time points 0 and 1, respectively. The following two equations were then applied to each biological replicate and the final doubling times averaged for each strain.

$$\text{Generation Time (Gt)} = \frac{\log_{10}(\text{Time 1}) - \log_{10}(\text{Time 0})}{\log_{10}(2)}$$

$$\text{Doubling Time (Dt)} = (\text{Gt} \times \Delta\text{Time}) \times 60$$

Detection of Delta-Haemolysin. The delta-haemolysin assay was performed as previously described (273). Each strain was tested in biological triplicate and any inconsistencies were resolved by accepting the results seen in two of the three tests.

Reverse Transcription - qPCR. Total RNA was extracted from cells grown to an OD600 of 1.0-1.1 in heart infusion broth using the RNeasy Mini Kit (Qiagen), as per the manufacturer's recommendations with the inclusion of two on-column DNase I digestion steps. Synthesis of cDNA and qPCR were performed as previously described (305), using RNAlII primers from (165). Control strain JKD6159 was used to establish a standard curve, and the average Ct from three technical replicates was used to calculate a fold difference for each strain compared to JKD6159. The average of two biological replicates is displayed in FIG 4A.

Murine Sepsis Model. Wild-type 8-week-old female BALB/c mice were injected via the tail vein with 100 μL of a phosphate buffered saline (PBS) solution containing approximate 1×10^6 colony forming units (CFU) of either BPH2032, BPH2057, BPH2070, BPH2073, BPH2080, BPH2086, BPH2095 or BPH2098. Mice were monitored every eight hours for signs of illness. Any mice that developed severe illness were euthanized by CO₂ inhalation. All surviving mice were euthanized at day seven post-injection. Bacterial burden within the liver, kidney and spleen were determined by tissue homogenisation in 5ml PBS and serial dilution. The minimum detectable CFU count was 5×10^4 CFU/mL. All animal work was performed in accordance with the Animal Research Ethics Committee at the University of Melbourne.

Statistical Analyses

All statistical analyses were performed in either R v2.15.3 (319), or Prism GraphPad v6.04 (320). Comparison of growth rates and MIC values were conducted in R using a Welsh two sample t-test. Significance was determined as a p-values ≤ 0.05 . In both analyses, the results from the four strains belonging to the small clade (BPH2006, BPH2008, BPH2010 and BPH2011) were included with Group 2. A log transformation was applied to MIC values during comparison. Analysis of results from the *in vivo* sepsis model was completed using Prism Graphpad. Kaplan-Meier curves were established, and log-rank tests used to compare survival. Differences in percentage weight loss from pre-injection weight was completed using either an unpaired t-test for comparison of year groups or a one-way ANOVA for RNA III expression groups. Significance was determined as a p value ≤ 0.05 .

All images were compiled and finalised using Inkscape v0.48.5 (321).

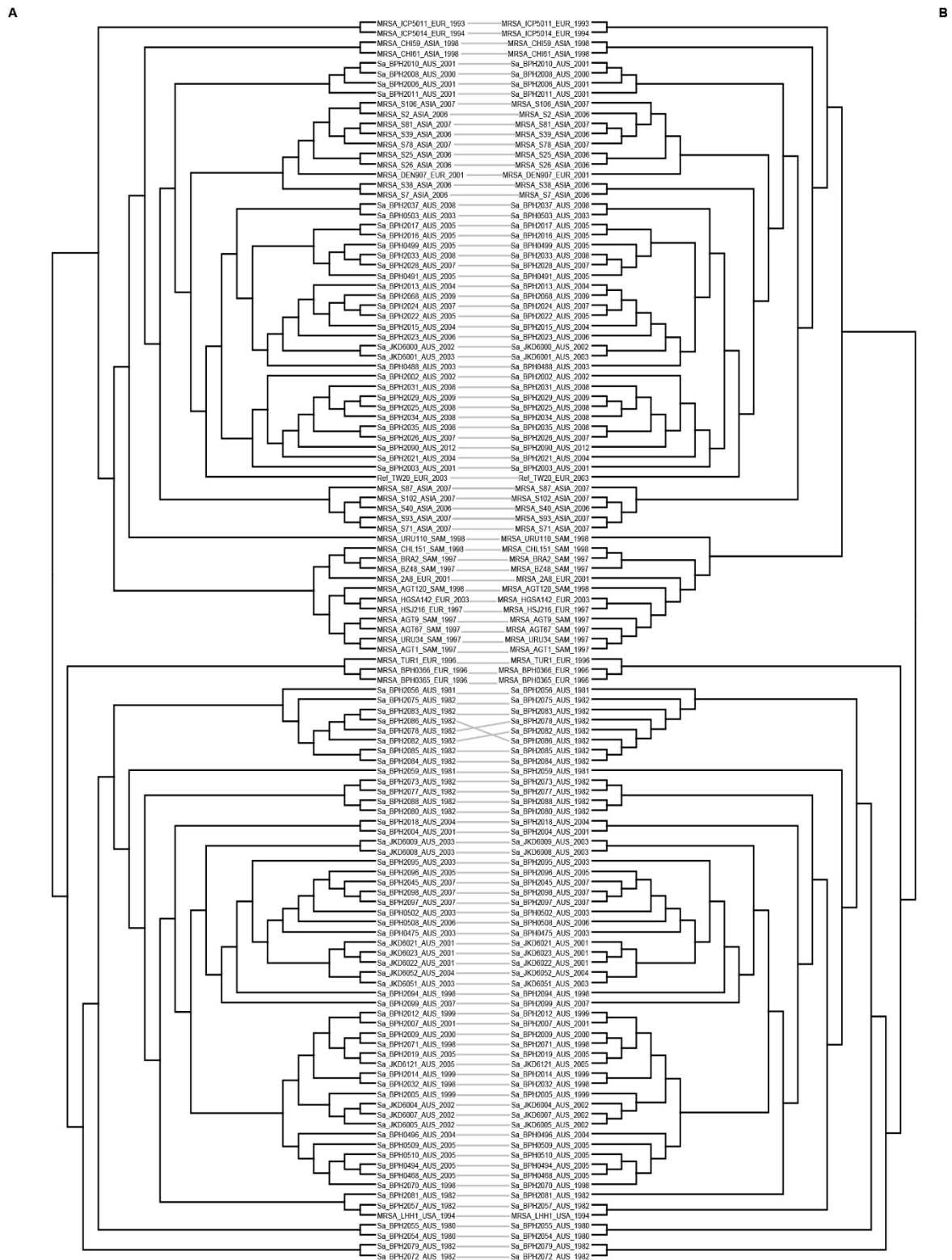


FIG S1 Comparison of the global phylogeny of ST239 MRSA inferred by Bayesian and ML methods. This tanglegram, generated in Dendroscope v3.2.10, compares the topology of the ST239 MRSA global phylogeny inferred by Bayesian (A) and ML (B) analyses. Both trees are displayed as cladograms. Branch annotations provide isolate names, continents, and years recovered and are joined by a grey line to indicate their location within the respective trees.

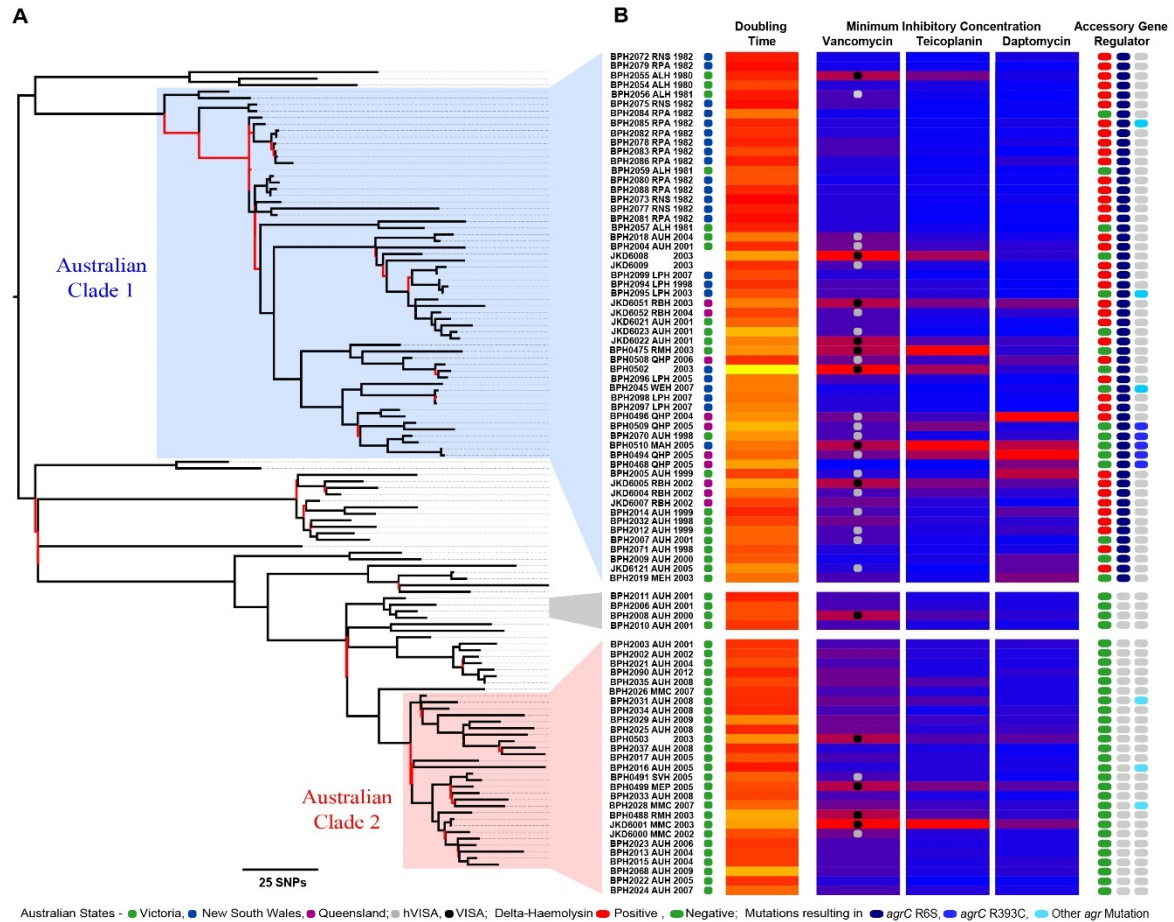


FIG S2 Phenotypic characterization of Australian ST239 *S. aureus* strains. (A) Global ST239 phylogeny presented in FIG 1A. Phenotypic data are presented for Australian isolates only, grouped into clades and ordered as they appear in the global phylogeny. Each isolate name is followed by the institution (three-letter code) at which it was isolated and then the year in which it was recovered. The Australian state is indicated by the coloured box (see key). (B) Heat maps representing the isolate specific phenotypes for doubling time (minutes; gradient of red [32.3] to yellow [76.1]) and MICs ($\mu\text{g/ml}$; gradient of blue to red) of vancomycin (0.5 to 4.0), teicoplanin (0.25 to 12.0), and daptomycin (0.094 to 4.0). Reduced glycopeptide susceptibility is indicated by a grey (hVISA) or black (VISA) circle. On the right, the function of the agr system, as determined by δ -hemolysin production, is shown. Mutations within the locus are displayed adjacent to the heat map (see key).

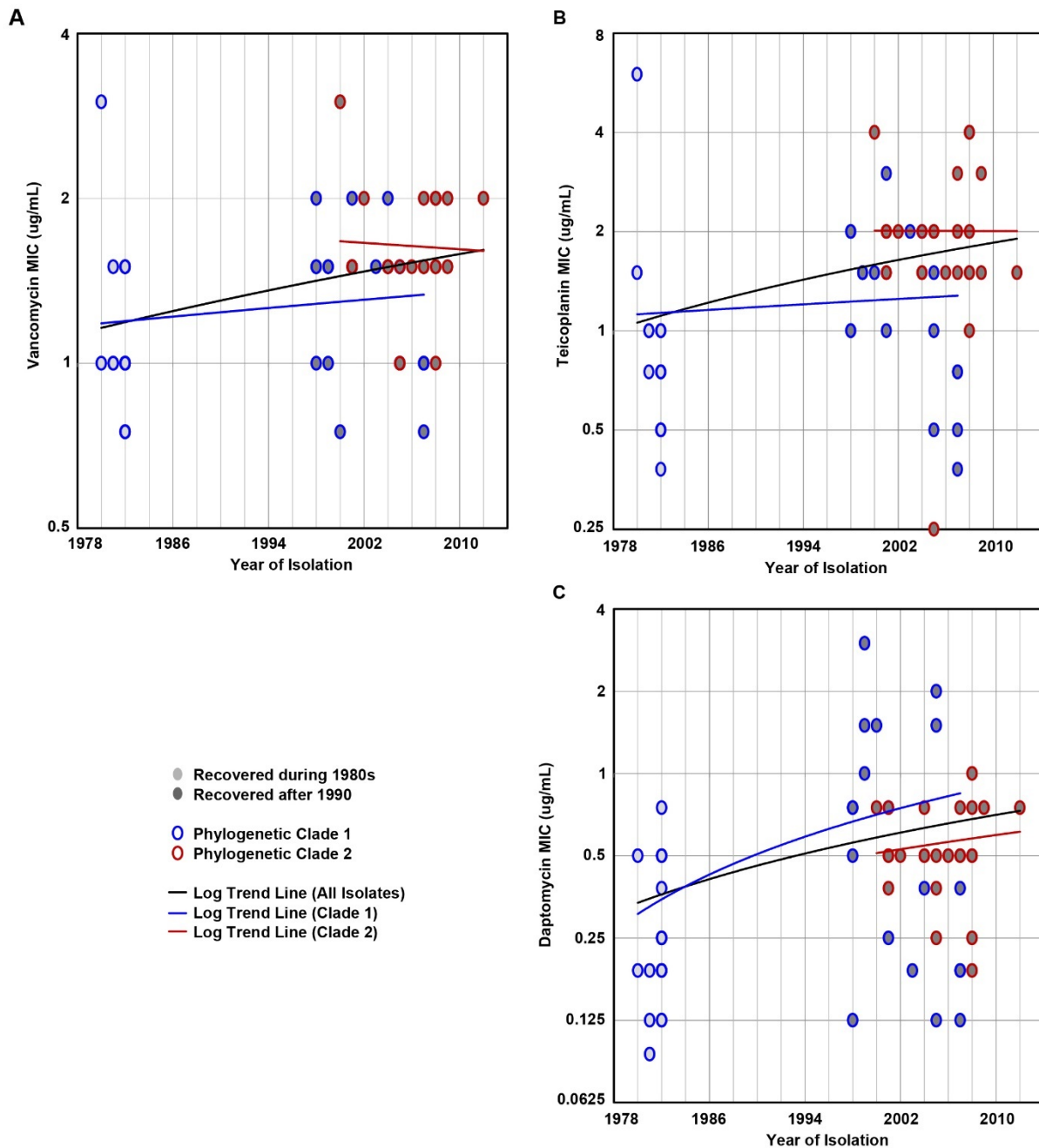


FIG S3 Increasing antimicrobial resistance of ST239 MRSA. The graphs illustrate changes in the MICs of vancomycin (A), teicoplanin (B), and daptomycin (C) for Australian ST239 MRSA over time. MICs are graphed on a log₂ scale. The temporal and phylogenetic clade of each isolate is indicated by the inner and outline colours of the data point, respectively (see key). Log trend lines used to estimate the MIC changes over time across the whole data set (black), Clade 1 (blue), and Clade 2 (red) are included.

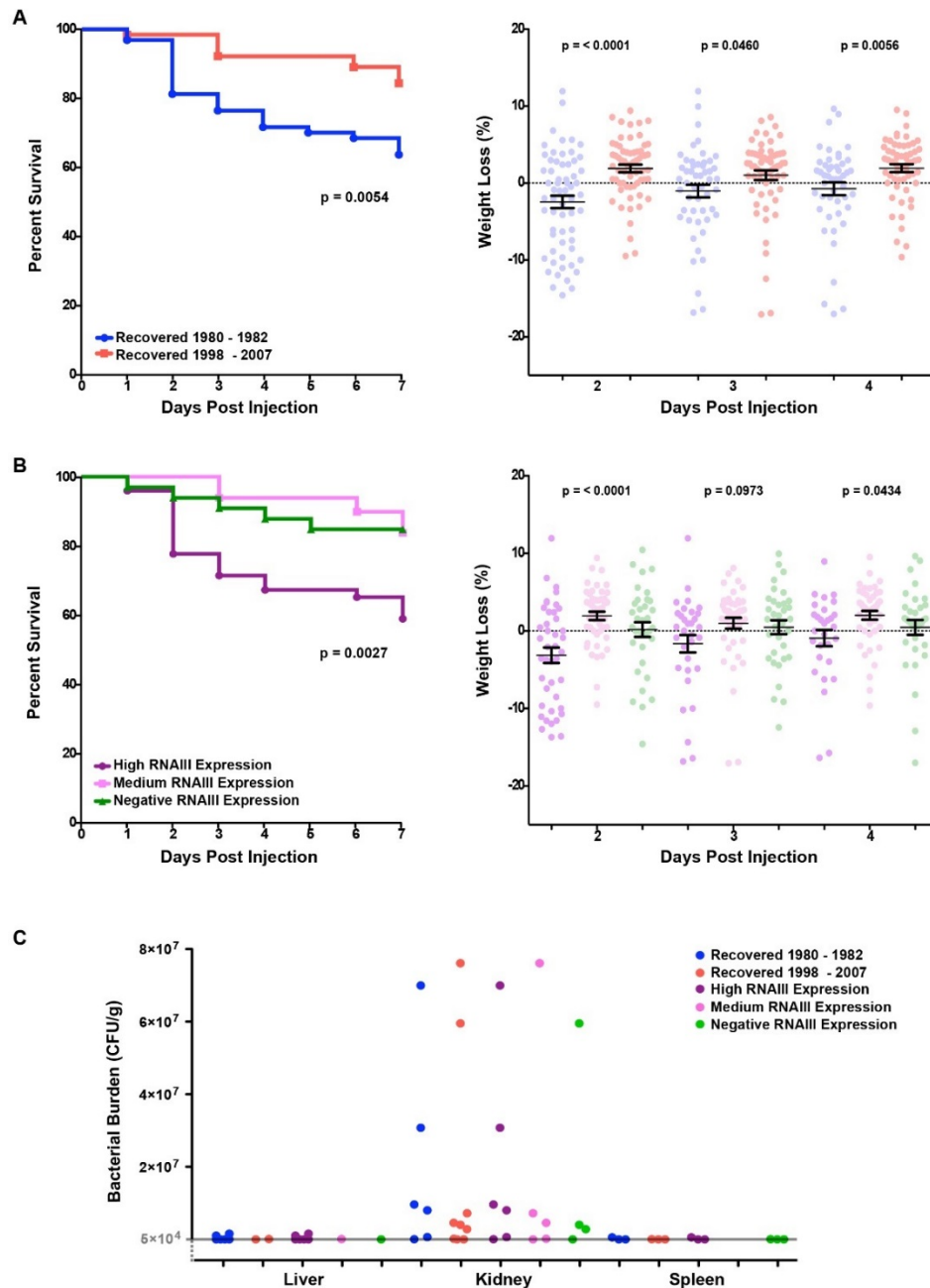


FIG S4 Evolution of ST239 *S. aureus* toward attenuated virulence. (A) Graphs showing the results of *in vivo* sepsis model analyses including 7-day survival curves and respective weight loss data for days 2, 3, and 4 post-injection, grouped by year of isolation. Each group represents four bacterial isolates, each tested in 16 animals. (B) Graphs illustrating the same data regrouped by RNAIII expression levels. The “high” and “medium” groups each represent three isolates, and the “low” group represents two isolates. In survival curves, P values are for comparisons of Kaplan-Meier curves with a log-rank test. In weight loss graphs, P values are for average weight loss comparisons of groups with an unpaired t test for year groups and one-way analysis of variance for RNAIII expression groups. Error bars show the standard errors of the means. (C) Bacterial burdens in the liver, kidneys, and spleen at 7 days post-injection. Each data point represents one animal infected with one isolate; however, each result is represented twice in the graph, once for the year group and once for each RNAIII expression group. The detection threshold is shown as a grey line; only values above this level are displayed.

Table S1: WalKR mutations in Australian ST239 MRSA

Mutation	Isolate Region Year	Clade	Vancomycin MIC (µg/mL) ^a	MET Interpretation
WalR _{A98T}	JKD6001 VIC 2003	2	4	VISA
WalR _{T103M}	BPH0503 VIC 2003	2	3	VISA
WalR _{K210R}	JKD6005 QLD 2002	1	3	VISA
	JKD6008 NZD 2003	1	4	VISA
WalK _{S9F}	BPH0488 VIC 2003	2	3	VISA
WalK _{M220I}	BPH2055 VIC 1982	1	3	VISA
WalK _{A243T}	BPH0510 NSW 2005	1	3	VISA
WalK _{V268F}	JKD6023 VIC 2001	1	1.5	hVISA
WalK _{A361V}	BPH2031 VIC 2008	2	2	VSSA
WalK _{M426I}	BPH0494 QLD 2005	1	2	hVISA
WalK _{A502V}	BPH0499 VIC 2005	2	3	VISA

^a As determined by Etest

MET, macro-Etest; MIC, minimum inhibitory concentration; VISA, vancomycin-intermediate *S. aureus*; hVISA, heterogeneous VISA; VSSA, vancomycin-susceptible *S. aureus*; VIC, Victoria; QLD, Queensland; NSW, New South Wales; NZD, New Zealand

Table S2 - Single nucleotide polymorphisms detected within the *agr* locus

Mutation (Location) ^a	Predicted Consequence	Prevalence Amongst Collection		
		<i>agr</i> Positive ^b	<i>agr</i> Negative ^b	Total
C-A (2172611)	AgrC _{R6S}	40	16	56
C-T (2173112)	AgrC _{P24S}	0	1	1
T-C (2173535)	AgrC _{F165L}	0	1	1
INS-C (2173719)	AgrC - Frame shift	1	0	1
G-A (2174192)	AgrC _{E384K}	0	1	1
C-T (2174219)	AgrC _{R393C}	0	5	5
T-C (2174754)	AgrA _{L134S}	0	1	1
INS-A (2175059)	AgrA - Frame shift	0	1	1
No mutation		0	27	27
No mutation (excluding AgrC _{R6S})		39	37	76

^aBase position in TW20

^b*agr* activity determined by the visual assessment of delta-haemolysis (present, positive; absent, negative)

2.3 Extended Analysis

One of the key findings of the study presented in section 2.2 was the discovery that two genetically distinct clades of ST239 circulated in Australia, the original clade (Clade 1) and a more recently introduced, genetically distinct clade (Clade 2), which from this point on will be referred to as the “Australian” and “Asian-Australian” clades, respectively. Due to the small number of isolates included in the publication, notably for the Asian-Australian clade, there was uncertainty surrounding the emergence date and geographic origin of the new clade. It was also unclear how far the Asian-Australian clone had spread within Australia. Since this work was conducted, the sequence data for more than 400 local and international ST239 isolates had become available. This extended analysis, utilising this new collection of 531 isolates, was performed to improve the phylogenetic modelling conducted in the original publication, and provide an updated description of ST239 MRSA in Australia.

2.3.1 Materials & Methods

Bacterial Isolates. This extended analysis utilised the genome sequence data for a total collection of 534 ST239 isolates. The source of these isolates and selection criteria are described below. Isolate specific information is available in Appendix 1. In the case of published isolates only the genome sequence data was utilised and sourced from public sequence archives. For all unpublished isolates, details have been provided about the storage, genomic DNA extraction and sequencing conditions.

Published sequence data

- The complete genomes of nine ST239 isolates were sourced from (216, 322-329).
- Seventy-nine isolates were sourced from the publication provided in section 2.2 (137). Of the original 86 isolates in this dataset, seven were excluded. These isolates had new sequence data generated as part of a separate study; new sequences have been included here as unpublished data.
- One hundred and seventy-seven isolates were sourced from the collections of (300, 301, 330-334). All ST239 isolates that could be accessed from these studies were included.
- Seventy-eight isolates were sourced from the collections of (31, 303). A limited number of isolates were included from both studies, selected to maximise the temporal and geographic diversity of these collections.

Unpublished sequence data

- One hundred and four isolates were sourced from the collection of the Australian and New Zealand Cooperative on Outcome in Staphylococcal Sepsis (ANZCOSS) study (335). All ST239 isolates from unique patients selected for an ANZCOSS sub-study were included (306).
- Twenty-nine isolates were sourced from the 2015 Australian Staphylococcal Sepsis Outcome Program (ASSOP). Sequence data for all ST239 isolates was kindly provided by the Australian Group on Antimicrobial Resistance (<http://agargroup.org.au/agar-surveys>).

- Sixteen isolates were sourced from the collection of the Vancomycin Efficacy in Staphylococcal Sepsis in Australasia (VANESSA) study (336). All ST239 isolates from unique patients were included.
- Forty-two isolates were sourced from historic collections (1989 – 1994) housed at the Microbiological Diagnostic Unit Public Health Laboratory (MDUPL), Melbourne, Australia. All ST239 isolates from unique patients that could be accessed were included.

Storage & Whole Genome Sequencing. Bacterial isolates sourced from the ANZCOSS, VANESSA and MDUPL collections were stored in glycerol broth at -80°C and sub-cultured twice on to horse blood agar (Oxoid). Genomic DNA was extracted with the JANUS Automated Workstation, using the Chemagic DNA/RNA kit (Perkin Elmer). Unique dual index libraries were prepared using the Nextera XT DNA preparation kit (Illumina). Libraries were sequenced on a NextSeq or HiSeq platform (Illumina) with 2 x 150 bp chemistry, or a MiSeq platform (Illumina) with 2 x 250 bp chemistry (Appendix 1), as per the manufacturer's instructions.

To generate a complete reference genome representative of the Asian-Australian lineage, isolate BPH2947 underwent long-read sequencing. Genomic DNA was extracted using the DNeasy Blood & Tissue extraction kit (Qiagen) and sequenced on the PacBio RS-II platform using P6-C4 chemistry, as per the manufacturer's instructions (Pacific Biosciences).

Genome Assembly of Reference BPH2947. The complete genome of BPH2947 was assembled using the SMRT Analysis System v2.3.0.140936 (Pacific Biosciences). Raw sequence data was *de novo* assembled using the HGAP v3 protocol, with a minimum seed read length of 10 Kb, genome size of 3 Mb, target coverage of 10 and an overlap error rate of 0.04. A single contig was generated. This contig underwent a second round of error correction using Quiver v1 and was circularised with the overlap removed in Geneious v8.1.5 (Biomatters). Raw long-reads were aligned to the assembly using BridgeMapper v1 and the alignment surrounding the contig overlap and ribosomal RNA regions were checked manually for assembly errors. The consensus sequence from this alignment underwent a final round of error correction, using Snippy v3.2-dev (<https://github.com/tseemann/snippy>), with short-read data generated on a MiSeq platform from the same genomic DNA extraction. The final sequence was annotated with Prokka v1.12 (337).

Phylogenetic Analysis. The Illumina short-read sequence data for all 534 ST239 isolates was mapped to reference genome BPH2947, using Snippy v3.2-dev. In the case of the nine complete genomes, pseudo short-reads (150 bp paired-end) were generated prior to mapping using the option "--ctgs" in Snippy v3.2-dev. A whole genome alignment was generated, using a minimum coverage of 10 reads and base-call stringency of 90% to call nucleotide variants. Examination of this alignment found three isolates to be highly divergent (data not shown). These isolates (GRE108, HU25, and TUR9) were excluded from further analysis. The whole genome alignment was reconstructed with the remaining 531 isolates. Recombination was assessed using ClonalFrameML v1.1.11 (338). However, as so few recombinant events were detected, masking these sites prior to temporal phylogenetic analysis was not required. Testing of nucleotide substitution models was performed on the whole genome alignment using IQ-TREE v1.6.1 (339), with the GTR nucleotide substitution model producing the highest likelihood.

The package BEAST v2.4.7 (340), was used to construct a time-aware phylogenetic model. The whole genome alignment was adjusted to create two smaller alignments, comprising only core genome SNPs; the core genome defined as a nucleotide site with a base called in every isolate within the alignment. These alignments represented: (i) All ST239 isolates (n=531, 7933 core SNPs), and (ii) all Australasian isolates (n=272, 3342 core SNPs). For each alignment, these core SNPs were grouped into a single partition and modelled using the following parameters. The year of isolation was provided as a prior. The GTR model of nucleotide substitution was selected, with the substitution rate estimated using four gamma categories. A UCLN clock model and a coalescent constant population model were selected [based on previous studies (137, 299)]. To model temporal changes in EPS in the Australian and Asian-Australian clades, the alignment of the Australasian isolates was split. Both alignments were modelled using the same parameters described above with the exception that the population model was set as a coalescent Bayesian skyline model. Invariant sites were included in all alignments as a constant pattern and weighted based on the nucleotide proportions of reference BPH2947 (A 33.4%, C 16.4%, G 16.4%, T 33.8%). Five 50 million step iterations were run for each alignment and then examined for convergence in Tracer v1.6 (<http://tree.bio.ed.ac.uk/software/tracer/>). The burnin was adjusted for each run to a point at which a stable likelihood was achieved. All runs converged and none were discarded based on the comparisons of Bayes Factors, calculated from the harmonic means of the run likelihoods, and the attainment of effective sample sizes (>200) in all measurements. Runs were joined and re-sampled at a frequency that would achieve 10,000 trees using LogCombiner. Bayesian skyline plots were generated in Tracer v1.6 using default parameters. MCC trees with median node heights were generated in TreeAnnotator and viewed in FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). All alterations to the aesthetics of the phylogenetic trees and Bayesian skyline plots were performed in either FigTree v1.4.3 or Inkscape v0.91 (<https://inkscape.org/en/>). Variation in the topology of the population model for all ST239 isolates was assessed for a sub-sample of 1,000 trees and visualised using DensiTree (341).

2.3.2 Results & Discussion

To re-examine the population structure of ST239 MRSA in Australia, the collection examined in the original publication (137), was supplemented with the sequence data for an additional 408 ST239 isolates. This collection of 531 isolates represented a time span of 35 years (1980 to 2015) and 31 countries across nine geographic regions: Africa (n = 2, Egypt and the Gambia), Australasia (n = 272, Australia and New Zealand), East Asia (n = 12, China and Taiwan), South-East Asia (n = 113, Malaysia, Singapore, Thailand and Vietnam), South Asia (n = 2, India and Sri Lanka), West Asia (n = 70, Syria and Turkey), Europe (n = 38, Czech Republic, Denmark, England, Germany, Greece, Hungary, Lithuania, Poland, Portugal, Romania and Spain), North America (n = 5, Canada and the United States), and South America (n = 17, Argentina, Brazil, Chile and Uruguay).

There has been strong agreement amongst population genomic studies that the global ST239 MRSA population structure is one of occasional intercontinental transmission, which has seeded this clone into new environments, followed by extensive local expansion, often displacing other HA-MRSA clones in the region (137, 300-302). The MCC tree, inferred from the 7,933 core genome SNPs of the 531 ST239 isolates was in agreement with this model and demonstrated a highly similar population structure (Figure 2.1). The large majority of isolates fell within one of six distinct clades (termed “lineages” in Figure 2.1): the European, South America, West-Asian, South-East Asian, and the Australian and Asian-Australian lineages. These represented the regional clades of ST239, each having arisen from a successful intercontinental transmission event and, for the most part, each clade represented only a single geographic region.

The inclusion of an additional 185 ST239 isolates from Australia, predominantly sourced from national cohorts, corroborates the findings of the original study in that there are three genetically distinct clades of Australian ST239 isolates, representing two lineages. These are the Australian lineage, which represented the original ST239 clone in Australia, and the Asian-Australian lineage, which represented two recent and independent introductions of the South-East Asian lineage into the Australian state of Victoria, with only one of these events proceeded by extensive local expansion (Figure 2.1 G, I). From this extended dataset, improved temporal estimates for the emergence dates of these lineages were obtained. The most recent common ancestor (MRCA) of the Australian lineage was estimated at 1971.11 (95% HPD range of 1967.41 to 1974.75) (Figure 2.2), a date consistent with early reports of ST239 in Australia (130, 296, 342). Whereas the MRCA of the Asian-Australian lineage was 1996.80 (95% HPD range of 1995.69 to 1998.68) for the larger clade (Figure 2.3 I). Both Asian-Australian clades shared core-genome similarity with isolates that had been recovered from Singapore (Figure 2.3). There was strong posterior support for the clustering of the larger Asian-Australian clade and two South-East Asian clades, that represented isolates recovered predominantly from healthcare institutions in Singapore between 2000-2010 (designated by point “H” in Figure 2.1, 2.3 and 2.4). Based on this analysis and the currently available sequence data, Singapore is the probable geographic origin for the two introductions of the Asian-Australian into Australia. There was less certainty around the geographic origin of the Australian lineage as there were no isolates in the collection recovered prior to 1980. This lineage appeared to be related, although distantly, to isolates that had been recovered from the United States (Figure 2.2). Further phylogeographic modelling of this clade is required.

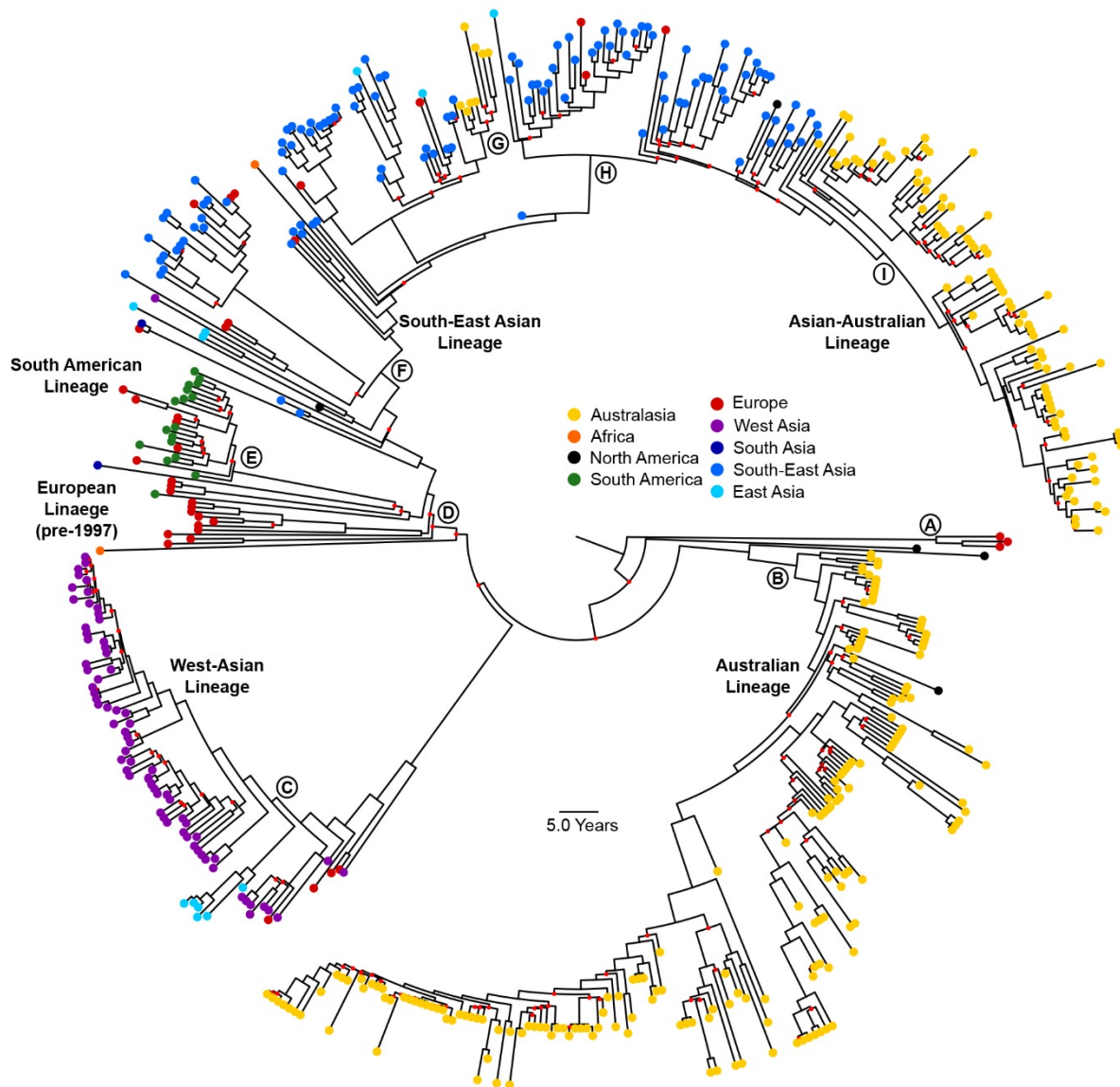


Figure 2.1 Global phylogeny of ST239 MRSA. Maximum clade credibility tree inferred from the 7,933 core genome SNPs of 531 ST239 isolates. Nodes with < 95% posterior support are indicated by a red circle. Geographic region of isolation for each isolate is indicated by the branch tip colour (refer to key). Nodes for regional clades are labelled: (A) divergent European clade (Greece), (B) Australian lineage, (C) West-Asian lineage (Turkey), (D) European lineage (pre-1997), (E) South American lineage, (F) South-East Asian lineage, (G) Asian-Australian lineage (small clade), (H) a clade of the South-East Asian lineage (Singapore) from which (I) the Asian-Australian lineage (large clade) has emerged. These regional clade labels are conserved across Figures 2.2, 2.3 and 2.4. Note, this figure is amended from FIG 1 of the publication presented in Chapter 4 (204).

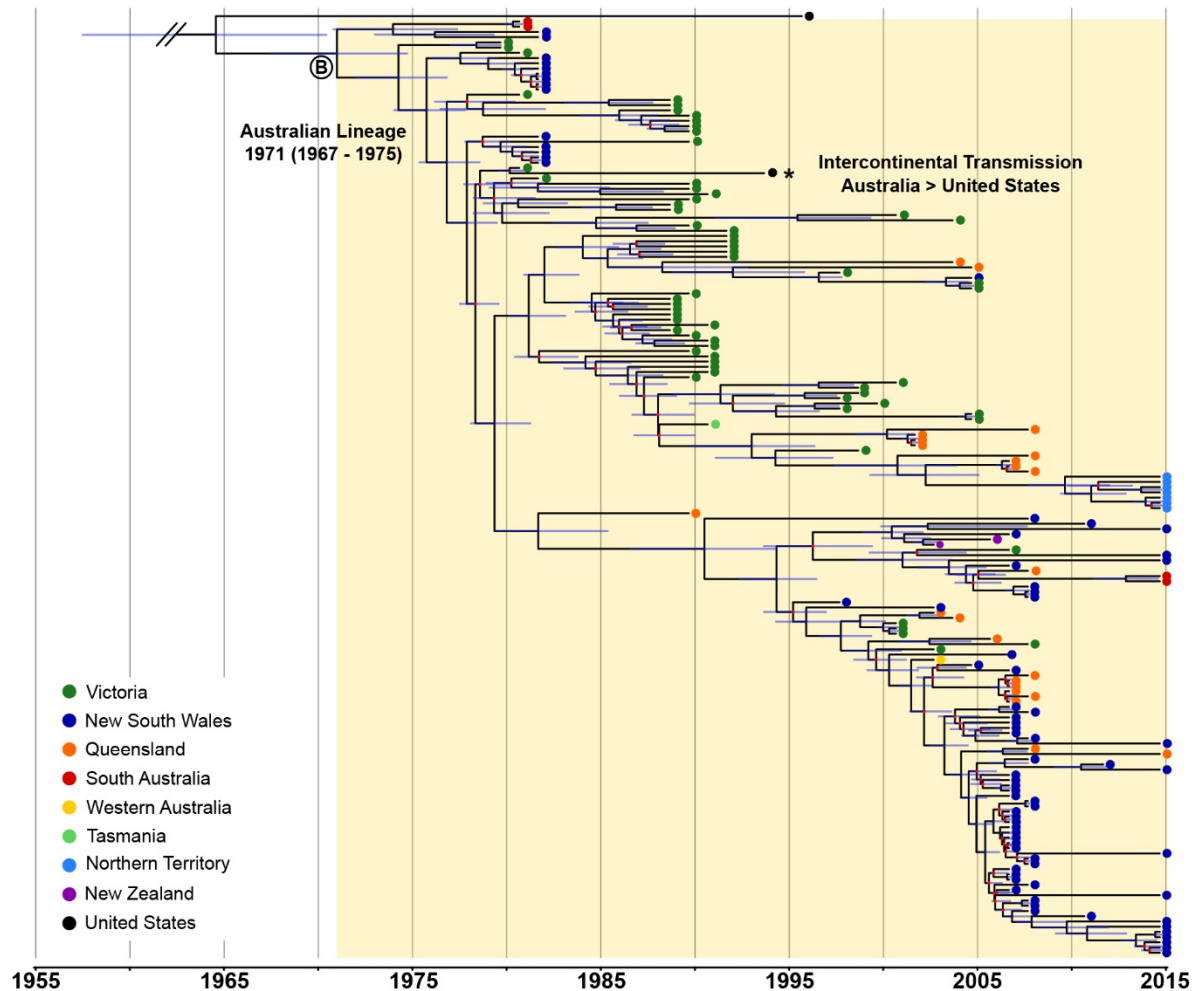


Figure 2.2 Population structure of the Australasian lineage. Maximum clade credibility tree is from Figure 2.1, with only the Australasian lineage displayed (yellow box). Nodes with < 95% posterior support are indicated by a red circle. Blue node bars illustrate the 95% HPD for the node height. Geographic region of isolation for each isolate is indicated by the branch tip colour (refer to key). The age and 95% HPD for the MRCA of the Australian lineage at node (B) is provided. An intercontinental transmission event of the Australian lineage into the United States is indicated.

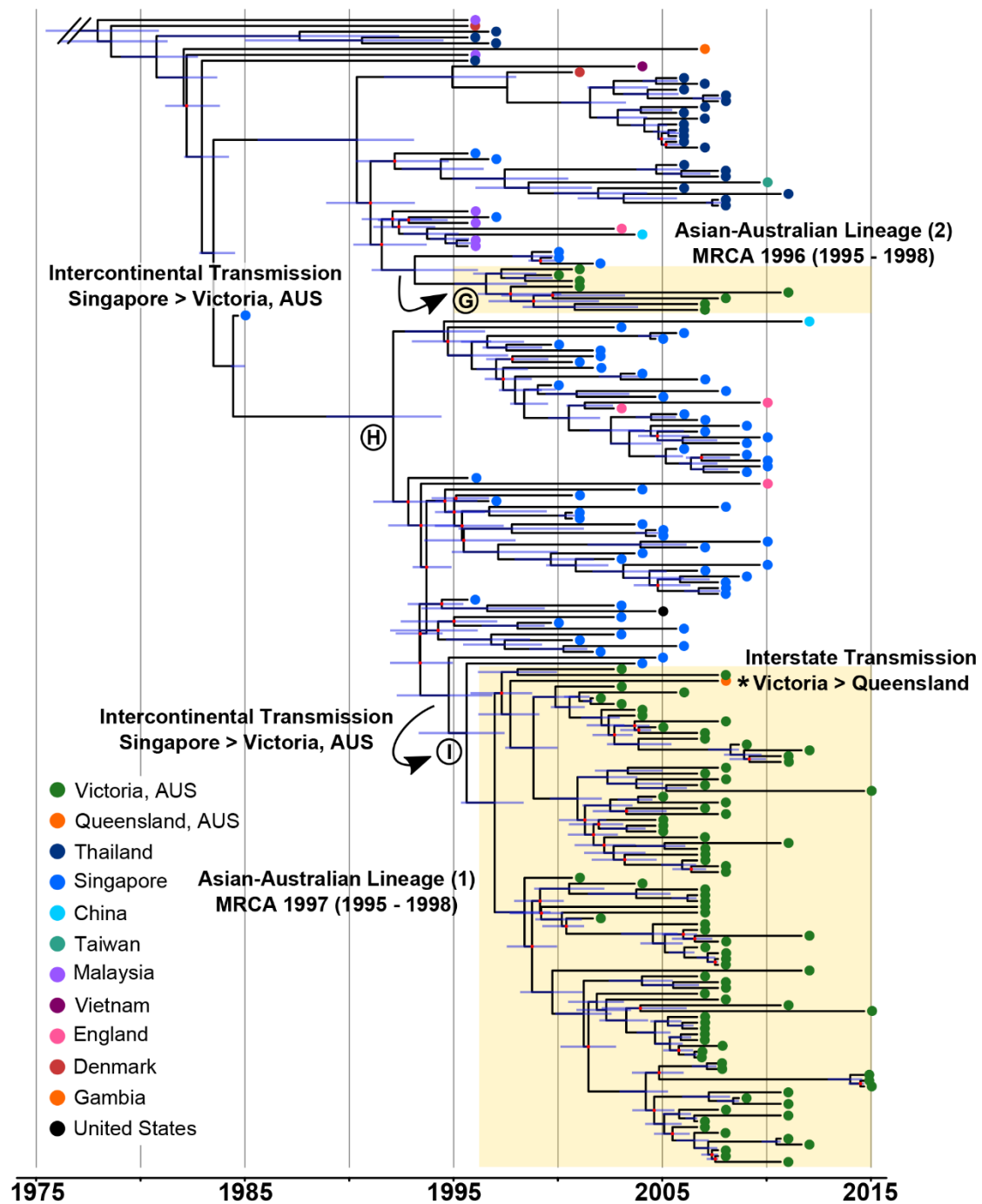


Figure 2.3 Population structure of the Asian-Australasian lineage. Maximum clade credibility tree is from Figure 2.1, with the Asian-Australasian lineage (yellow box) and closely related South-East Asian clades displayed. Nodes with < 95% posterior support are indicated by a red circle. Blue node bars illustrate the 95% HPD for the node height. Geographic region of isolation for each isolate is indicated by the branch tip colour (refer to key). The age and 95% HPD for the MRCA of the two Asian-Australian clades at node (G) and (I) is provided. Interstate transmission of the Asian-Australasian lineage from the state of Victoria into Queensland is indicated.

As of 2015, the Australian lineage continued to circulate in the Australian states of New South Wales, South Australia, Queensland and the Northern Territories (Figure 2.2). The Asian-Australian lineage remained restricted to the state of Victoria, with only a single interstate transmission event into Queensland detected (Figure 2.3). The geographic restriction of this lineage may in part be due to the autonomy of the healthcare services in the different Australian states, with the rarity of interstate patient transfer minimising the opportunity for interstate movement of ST239 MRSA. Furthermore, the prevalence of ST239 amongst HA-MRSA has decreased over the last two decades, from 86.8% in 2003 to only 21.2 % in 2017 (<http://agargroup.org.au/agar-surveys>). This decrease in prevalence was also observed in the EPS modelling for the Australian and Asian-Australian lineages (Figure 2.5). The temporal EPS trend for the Australian lineage demonstrated a clear peak in population genetic diversity around the late 1980s and early 1990s, which suggested population expansion during this time. However, from the 1990s the EPS decreased and then plateaued. The introduction of the Asian-Australian lineage in the mid-1990s boosted the population during the early 2000s but it has also since plateaued (Figure 2.5). These patterns suggested that the genetic diversity of the ST239 population is currently stable. Although the data from national surveillance programs indicates that the prevalence of ST239 is still in decline, whilst the total incidence of MRSA has decreased to a lesser extent (<http://agargroup.org.au/agar-surveys>). This discrepancy between the EPS models and national surveillance data may reflect the fact that the contemporary ST239 population in this collection (those isolates recovered between 2010 – 2015) represented the same genetic diversity as the population from the preceding decade (2000 – 2010) (Figure 2.2), simply in reduced numbers. If this population decrease continues, the population genetic diversity will likely diminish further and ST239 MRSA may even disappear completely from Australian hospitals.

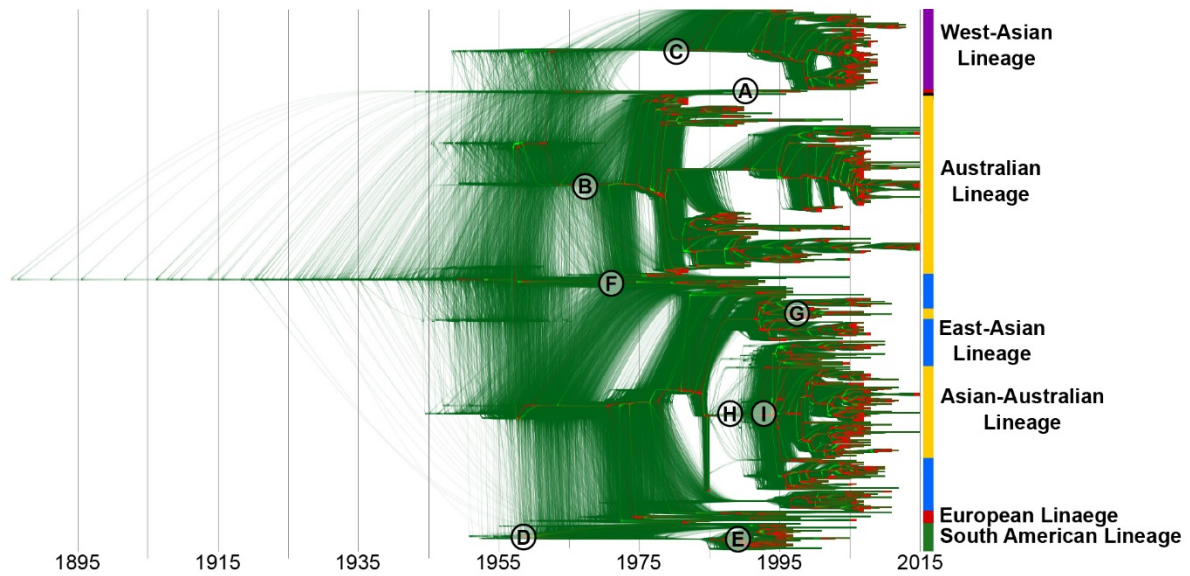


Figure 2.4 Topological variation in the global phylogeny of ST239 MRSA. Displayed is an overlay of 1000 trees sampled from the same Bayesian analysis used to generate MCC tree in Figure 2.1, visualised with Densitree. Blue, red and light green lines indicate the topologies with highest likelihoods (in that order). All other topologies are coloured dark green. Node labels are the same as those in Figure 2.1, indicating regional clades. Major lineages are annotated alongside the phylogeny for reference.

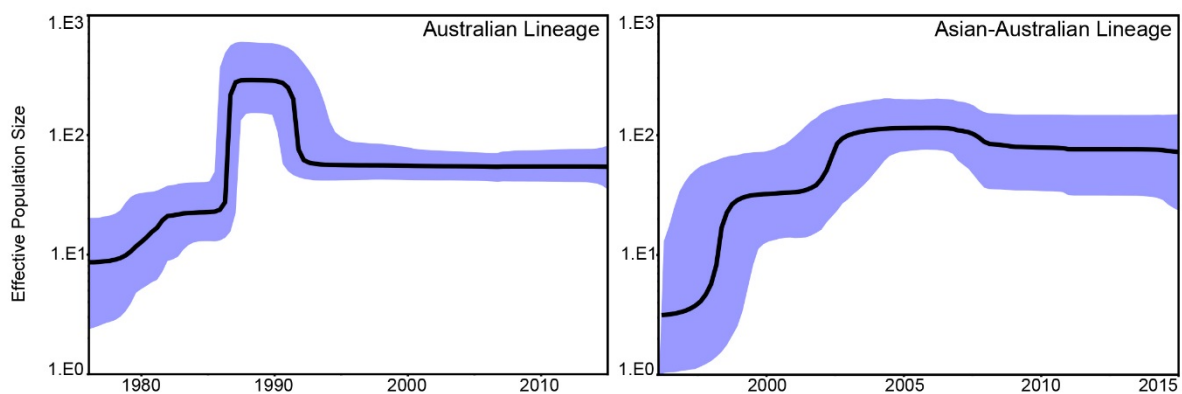


Figure 2.5 Fluctuations in effective population size overtime. Bayesian skyline plots, inferred from the 3,342 core genome SNPs of 272 Australasian ST239 isolates, illustrate fluctuations in the EPS over time in the Australian and Asian-Australian lineages. Central black line indicates the median effective population size, and the purple boundary indicates the 95% HPD.

2.4 Conclusion

Using a combination of population and functional genomics this work has provided insight into the local evolution of one of the most globally significant HA-MRSA lineages. Australia has not only one but two successful lineages of ST239 that co-circulate. The previously unrecognised importation of ST239 MRSA into Australia from South-East Asia in the late 1990s had contributed to the persistence of this multidrug resistant staphylococcal lineage as a cause of healthcare-associated infection. In the context of a geographically diverse dataset, it appeared that the Asian-Australian lineage originated in Singapore. Phenotypically it displayed enhanced antimicrobial resistance, in particular against key anti-MRSA antibiotics, and this has likely provided it with an evolutionary advantage within the healthcare environment. This highly selective environment also appears to have influenced the evolution of the original Australian lineage, favouring the selection of isolates within the population that demonstrate reduced antibiotic susceptibility. The strong selective pressures imposed by healthcare environments is evident and reinforced by the convergent evolution in these two lineages. The apparent cost of adaptive evolution in ST239 towards increased antimicrobial resistance appears to be a significant reduction in virulence potential and replicative fitness. Whether this is a true evolutionary consequence or co-selected adaptation remains unclear, as the capacity of ST239 to cause persistent infection remains significant. Together these results reveal the complexity of the evolutionary dynamics surrounding MRSA lineages that predominate within healthcare environments and illustrates the benefit of utilising a combination of genomic and phenotypic investigation to enhance our understanding of the mechanisms and environmental factors influencing the adaptation of microbial populations.

Chapter 3

Rapid emergence and evolution of *Staphylococcus aureus* clones harbouring *fusC*-containing staphylococcal cassette chromosome elements

3.1 Introduction

Skin and soft tissue infection is the most common clinical manifestation of *S. aureus* disease, and reports from multiple regions suggest that the incidence of SSTI is increasing (34-37). New Zealand has one of the highest rates of SSTI in the industrial world, and *S. aureus* is the leading bacterial cause of SSTI-associated morbidity and mortality in the region (37, 343). Methicillin-susceptible *S. aureus* (MSSA) is responsible for the great majority of staphylococcal disease in New Zealand, with rates of MRSA having remained stable at approximately 8-10% (https://surv.esr.cri.nz/antimicrobial/staphylococcus_aureus.php).

The majority of SSTIs occur in the community and treatment may involve the use of a topical antibiotic (344); the use of either topical fusidic acid (FA) or mupirocin found to be superior to a placebo in the context of mild to moderate impetigo (179). FA is both a topical and systemic antibiotic active against many Gram-positive bacterial species and the most commonly prescribed topical antibiotic in New Zealand (39). However, its unrestricted use in the community has seen resistance rates rise amongst *S. aureus*, from 17% in 1999 to 25-28% in 2013-2014 (39). This increase in resistance was associated with the emergence of three FA-resistant clones: ST5 MRSA (colloquially known as “AK3”), a ST1 MRSA (“AK1”) and a ST1 MSSA (345). In all three clones, resistance is mediated by the presence of *fusC*, an acquired resistance gene that is invariably carried by SCC elements and has been identified in *S. aureus* isolates recovered from multiple regions (47, 193, 346, 347).

Prior to this work the origin of these FA-resistant clones in New Zealand and the genetic context and role of *fusC* in their emergence had not been explored. It was unclear if the high population level use of FA had created an environment that promoted the local development of these resistant lineages or provided an opportunity for an imported clone to thrive, where FA resistance provided an adaptive advantage over local circulating clones. Understanding the evolutionary dynamics surrounding the recent and rapid emergence of these *S. aureus* clones in the community environment is crucial in informing decisions regarding the regulation of topical FA to reduce the development and further spread of resistance.

3.1.1 Project Aims

The previous chapter explored the evolutionary dynamics of a globally dispersed clone of MRSA that has become well adapted to the healthcare environment. However, MRSA is also a growing problem in the community setting. The primary aim of the work presented in this chapter was to use population genomics to explore the evolutionary dynamics surrounding the recent and very rapid emergence of a new multiply antibiotic resistant MRSA clone in the New Zealand community. This included: (i) generating a population model for FA-resistant ST5 MRSA, both locally and in the context of a globally diverse dataset, in order to gain a deeper understanding of the geographic origin and local expansion of this clone; (ii) exploring the genetic context surrounding *fusC* in FA-resistant MRSA in New Zealand. This includes a comparison to what has been observed internationally in order to further investigate the role of the SCC in the global dissemination of *fusC* and provide insight into the consequences surrounding the emergence and spread of FA-resistant *S. aureus* in populations with high-level and unregulated FA use.

3.1.2 Chapter Overview

Presented in this chapter are the results of experiments conducted to address the two objectives described above, both of which have been answered in the publication presented in section 3.2. The supplementary figure from this publication is presented in section 3.3, with the supplementary dataset provided in Appendix 2. In addition to this publication, an extended genomic analysis is presented in section 3.3. This extended analysis makes use of the sequence data for an additional 376 CC5 isolates that had recently become available to revisit the emergence of FA-resistant ST5 MRSA in New Zealand in the context of a large international dataset. A final summary of this chapter is presented in section 3.4. The implications and limitations of this work within the wider context of this thesis are discussed in Chapter 5.

3.2 Results

3.2.1 Publication: Rapid emergence and evolution of *Staphylococcus aureus* clones harboring *fusC*-containing staphylococcal cassette chromosome elements

Baines, S. L., B. P. Howden, H. Heffernan, T. P. Stinear, G. P. Carter, T. Seemann, J. C. Kwong, S. R. Ritchie and D. A. Williamson (2016). Antimicrobial Agents and Chemotherapy. DOI: 10.1128/AAC.03020-15.

Rapid Emergence and Evolution of *Staphylococcus aureus* Clones Harboring *fusC*-Containing Staphylococcal Cassette Chromosome Elements

Sarah L. Baines,^{a,b} Benjamin P. Howden,^{a,b,c} Helen Heffernan,^d Timothy P. Stinear,^{a,b} Glen P. Carter,^{a,b} Torsten Seemann,^{a,e} Jason C. Kwong,^{a,b,c} Stephen R. Ritchie,^f Deborah A. Williamson^{a,b,d}

Doherty Applied Microbial Genomics, Department of Microbiology & Immunology, The University of Melbourne at The Doherty Institute for Infection and Immunity, Melbourne, Australia^a; Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology & Immunology, The University of Melbourne at The Doherty Institute for Infection and Immunity, Melbourne, Australia^b; Infectious Diseases Department, Austin Health, Melbourne, Australia^c; Institute of Environmental Science and Research, Wellington, New Zealand^d; Victorian Life Sciences Computation Initiative, The University of Melbourne, Melbourne, Australia^e; School of Medical Sciences, University of Auckland, Auckland, New Zealand^f

The prevalence of fusidic acid (FA) resistance among *Staphylococcus aureus* strains in New Zealand (NZ) is among the highest reported globally, with a recent study describing a resistance rate of approximately 28%. Three FA-resistant *S. aureus* clones (ST5 MRSA, ST1 MSSA, and ST1 MRSA) have emerged over the past decade and now predominate in NZ, and in all three clones FA resistance is mediated by the *fusC* gene. In particular, ST5 MRSA has rapidly become the dominant MRSA clone in NZ, although the origin of FA-resistant ST5 MRSA has not been explored, and the genetic context of *fusC* in FA-resistant NZ isolates is unknown. To better understand the rapid emergence of FA-resistant *S. aureus*, we used population-based comparative genomics to characterize a collection of FA-resistant and FA-susceptible isolates from NZ. FA-resistant NZ ST5 MRSA displayed minimal genetic diversity and represented a phylogenetically distinct clade within a global population model of clonal complex 5 (CC5) *S. aureus*. In all lineages, *fusC* was invariably located within staphylococcal cassette chromosome (SCC) elements, suggesting that SCC-mediated horizontal transfer is the primary mechanism of *fusC* dissemination. The genotypic association of *fusC* with *mecA* has important implications for the emergence of MRSA clones in populations with high usage of fusidic acid. In addition, we found that *fusC* was colocated with a recently described virulence factor (*tirS*) in dominant NZ *S. aureus* clones, suggesting a fitness advantage. This study points to the likely molecular mechanisms responsible for the successful emergence and spread of FA-resistant *S. aureus*.

The commonest clinical manifestation of *Staphylococcus aureus* disease is skin and soft-tissue infection (SSTI). In most cases, *S. aureus* SSTI is diagnosed and managed in the primary care setting, and depending on the clinical manifestation, treatment may involve the administration of a topical antimicrobial agent, such as fusidic acid (FA) (1). FA is a topical and systemic antimicrobial and is active against several Gram-positive organisms. Recently licensed in the United States, FA has been used extensively in several settings for the topical treatment of superficial *S. aureus* SSTI (2). We have recently described a significant increase in topical FA use in New Zealand (NZ), concurrent with an increase in FA resistance rates in *S. aureus*, from 17% in 1999 to 28% in 2013 (2). This increase in resistance was associated with the emergence of three specific *S. aureus* clones: multilocus sequence type 5 (ST5) methicillin-resistant *S. aureus* (MRSA), ST1 methicillin-susceptible *S. aureus* (MSSA), and ST1 MRSA. In particular, an ST5 MRSA clone (colloquially known as AK3 ST5 MRSA) has emerged over the past decade to become the dominant community- and hospital-associated MRSA clone in NZ (3). First reported in 2005, AK3 ST5 MRSA strains harbor a type IV staphylococcal cassette chromosome (SCC) *mec* and, in addition to being resistant to penicillin and oxacillin, generally are resistant to FA, with one recent study describing an FA resistance rate of 89% in this clone (2). To date, however, the origin and evolutionary history of AK3 ST5 MRSA has not been determined, and it is unknown whether this clone has emerged from locally circulating ST5 MSSA or, alternatively, has been imported into NZ with subsequent regional transmission.

In all three clones (AK3 ST5 MRSA, ST1 MSSA, and ST1 MRSA), FA resistance is associated with the presence of the *fusC* gene (2). However, the origin and genetic context of the *fusC* resistance locus has not been investigated previously in *S. aureus* from NZ. To date, only a few studies have previously described the genetic context of *fusC* among small numbers of *S. aureus* lineages (4–7). Furthermore, the extent to which common mobile genetic elements are shared among *fusC*-harboring *S. aureus* clones globally has not been investigated previously but could provide important insights into the mechanisms of emerging resistance to topical antimicrobials across different *S. aureus* lineages.

Received 19 December 2015 Returned for modification 7 January 2016

Accepted 22 January 2016

Accepted manuscript posted online 8 February 2016

Citation Baines SL, Howden BP, Heffernan H, Stinear TP, Carter GP, Seemann T, Kwong JC, Ritchie SR, Williamson DA. 2016. Rapid emergence and evolution of *Staphylococcus aureus* clones harboring *fusC*-containing staphylococcal cassette chromosome elements. *Antimicrob Agents Chemother* 60:2359–2365. doi:10.1128/AAC.03020-15.

Address correspondence to Deborah A. Williamson, deborah.williamson@unimelb.edu.au.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AAC.03020-15>.

Copyright © 2016 Baines et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Accordingly, the objectives of this study were to (i) investigate the evolution of the AK3 ST5 MRSA clone in NZ, (ii) characterize the genetic context of the *fusC* gene in dominant FA-resistant clones in NZ, namely, AK3 ST5 MRSA, ST1 MSSA, and ST1 MRSA, and (iii) compare these to international *fusC*-harboring isolates of *S. aureus* to better understand the rapid emergence of FA-resistant *S. aureus*.

MATERIALS AND METHODS

Bacterial isolates. A total of 46 isolates of *S. aureus* recovered in NZ were used in this study. These isolates were sourced from (i) national surveys of MRSA carried out between 2005 and 2013 ($n = 34$) (https://surv.esr.cri.nz/antimicrobial/mrsa_annual.php) and (ii) a contemporary molecular epidemiological survey of *S. aureus* carried out in Auckland, NZ, in 2013 ($n = 12$) (2). Between two and six representatives of clonal complex 5 (CC5) were selected from each survey year, and six representatives of both ST5 and ST1 were selected from the contemporary molecular epidemiological survey. This collection additionally was supplemented with the fully assembled genomes of 11 *S. aureus* strains belonging to CC5 (8–16), the CC1 strain MSSA476 (4), and the genome sequence reads of 12 FA-resistant *S. aureus* strains from a recent study in the United Kingdom (UK) (6). All CC5 *S. aureus* strains from this study were included, as well as a single representative for each SCC element described for other *S. aureus* lineages. Further information for each isolate is available in Table S1 in the supplemental material.

Genome sequencing. Genomic DNA was extracted using the DNeasy blood and tissue extraction kit (Qiagen), and DNA libraries were created using the Nextera XT DNA preparation kit (Illumina). Sequencing was performed on either the HiSeq platform (Illumina) with 2- by 100-bp paired-end chemistry or the MiSeq platform (Illumina) with 2- by 250-bp paired-end chemistry. *S. aureus* isolate NZAK3 was sequenced on the Pacific Biosciences RS-II platform.

Genome assembly of NZAK3. The complete genome assembly of *S. aureus* reference strain NZAK3 was performed using the SMRT analysis system v2.3.0.140936 (Pacific Biosciences). Raw sequence data were *de novo* assembled using the HGAP3 protocol, with a minimum seed read length of 5,000, genome size of 3 Mb, target coverage of 10, and overlap error rate of 0.04. Polished contigs were further error corrected using Quiver v1. The read alignments were visually assessed, and one rRNA region was found to have poor support; subsequently, the contig was cut at this region. The resulting contigs then were ordered against *S. aureus* N315 (8) using Mauve (17) and joined in Geneious v8.1.5 (Biomatters); the altered assembly around the rRNA locus was confirmed with long-range PCR. The final assembly was checked using BridgeMapper v1 in the SMRT analysis system, and the consensus sequence was corrected with short-read Illumina data using Snippy v2.5 (<https://github.com/tseemann/snippy>). The final sequences were annotated using Prokka v1.11 (18), and *in silico* multilocus sequence typing (MLST) was performed using MLST v1.2 (<https://github.com/tseemann/mlst>). Sequence data for NZAK3 have been deposited in the European Nucleotide Archive (ENA) under primary accession PRJEB12333.

Phylogenetic analysis. The genomes of the 11 assembled CC5 *S. aureus* strains first were shredded into 100-bp pseudoreads with 50× coverage using Nesoni v0.130 (<https://github.com/Victorian-Bioinformatics-Consortium/nesoni>). The sequence reads of all CC5 *S. aureus* strains then were mapped to reference strain NZAK3, and single-nucleotide polymorphisms (SNPs) were detected using Snippy v2.5 with a minimum coverage depth of 10 and base call stringency of 90%. A maximum likelihood (ML) phylogenetic tree was constructed from core genome SNPs using PhyML v3.0 (19) run under a general time-reversible model with 1,000 bootstrap replicates.

Analysis of the genetic context of *fusC* and alignment of *fusC*-containing SCC elements. The genome sequence reads of all unassembled isolates were *de novo* assembled using SPAdes v3.6.1 (20). The program contig-puller (<https://github.com/kwongj/contig-puller>) then was used

to identify, extract, and annotate contigs that contained a *fusC* gene with 99% sequence identity to that present in *S. aureus* MSSA476 (GenBank accession number BX571857, locus tag SAS0043). These contigs were manually examined in Geneious v8.1.5 (Biomatters). Sequence alignments were performed using ClustalW v2.1 (21) and the Artemis comparison tool v16.0.11 (22).

BioProject sequence accession number. All sequence data are available at the ENA under primary accession no. PRJEB12333.

RESULTS AND DISCUSSION

Complete genome sequence of *S. aureus* NZAK3. To characterize AK3 ST5 MRSA and provide a closely related reference genome for subsequent phylogenetic analysis, the first reported strain of AK3 ST5 MRSA was sequenced and completely assembled using PacBio sequencing. This strain, subsequently referred to as NZAK3, was isolated in Auckland, NZ, in 2005 from a patient with SSTI. To our knowledge, this is the earliest isolation of this FA-resistant AK3 ST5 MRSA clone globally.

The genome of NZAK3 comprises a circular chromosome spanning 2,831,681 bp (GC content of 32.8%) and a single circular plasmid (pNZAK3) spanning 25,689 bp (GC content of 29.5%). A total of 2,607 protein-coding regions (CDS) were identified in the chromosome, and 30 CDS were identified in pNZAK3. The chromosome contained five rRNA operons and 62 tRNA genes. *In silico* MLST confirmed that this isolate belonged to ST5.

Comparison of the chromosome of NZAK3 to all other completely assembled *S. aureus* genomes belonging to CC5 ($n = 11$) illustrated a high level of sequence conservation among members of this lineage (Fig. 1). The accessory CC5 genome consisted primarily of mobile elements, specifically the SCCmec locus, a chromosomal copy of *blaZ*, and two complete prophages. These prophages were presumptively identified as phiETA2 (GenBank accession no. NC_008798) and phiN315 (GenBank accession no. NC_004740) based on a predominance of high-identity BLAST matches to each of these elements. The plasmid pNZAK3 was mostly unique, with only 45% (11,588/25,689 bp) of the sequence matching a previously published plasmid sequence with 99% DNA identity (plasmid SAP059A; GenBank accession no. GQ900480.1). The remaining region in pNZAK3 contained a cadmium resistance locus but no other known resistance or virulence determinants.

Population structure and evolution of AK3 ST5 MRSA. To characterize the population structure of the AK3 clone and investigate its evolutionary origins, 39 isolates of *S. aureus* recovered in NZ and belonging to CC5 underwent whole-genome sequencing. To maximize temporal and geographic diversity within the collection, between two and six isolates were selected from each year between 2006 and 2013 and from different regions in NZ (detailed in Materials and Methods). This collection also was supplemented with the 11 fully assembled CC5 genome sequences described above and the genome sequence reads of eight CC5 MRSA isolates from a recent study of FA-resistant *S. aureus* in the United Kingdom (6). Details of these isolates are available in Table S1 in the supplemental material.

Mapping of the genome sequence data of these 58 CC5 *S. aureus* strains to reference strain NZAK3 identified a total of 14,584 SNPs, of which 4,233 were present within the core genome (87.3% of reference strain NZAK3). The 4,233 core genome SNPs were used to construct an ML phylogenetic tree (Fig. 2).

This collection of CC5 *S. aureus* isolates represented five sequence types: ST5, ST73, ST105, ST149, and ST835 (Fig. 2). Iso-

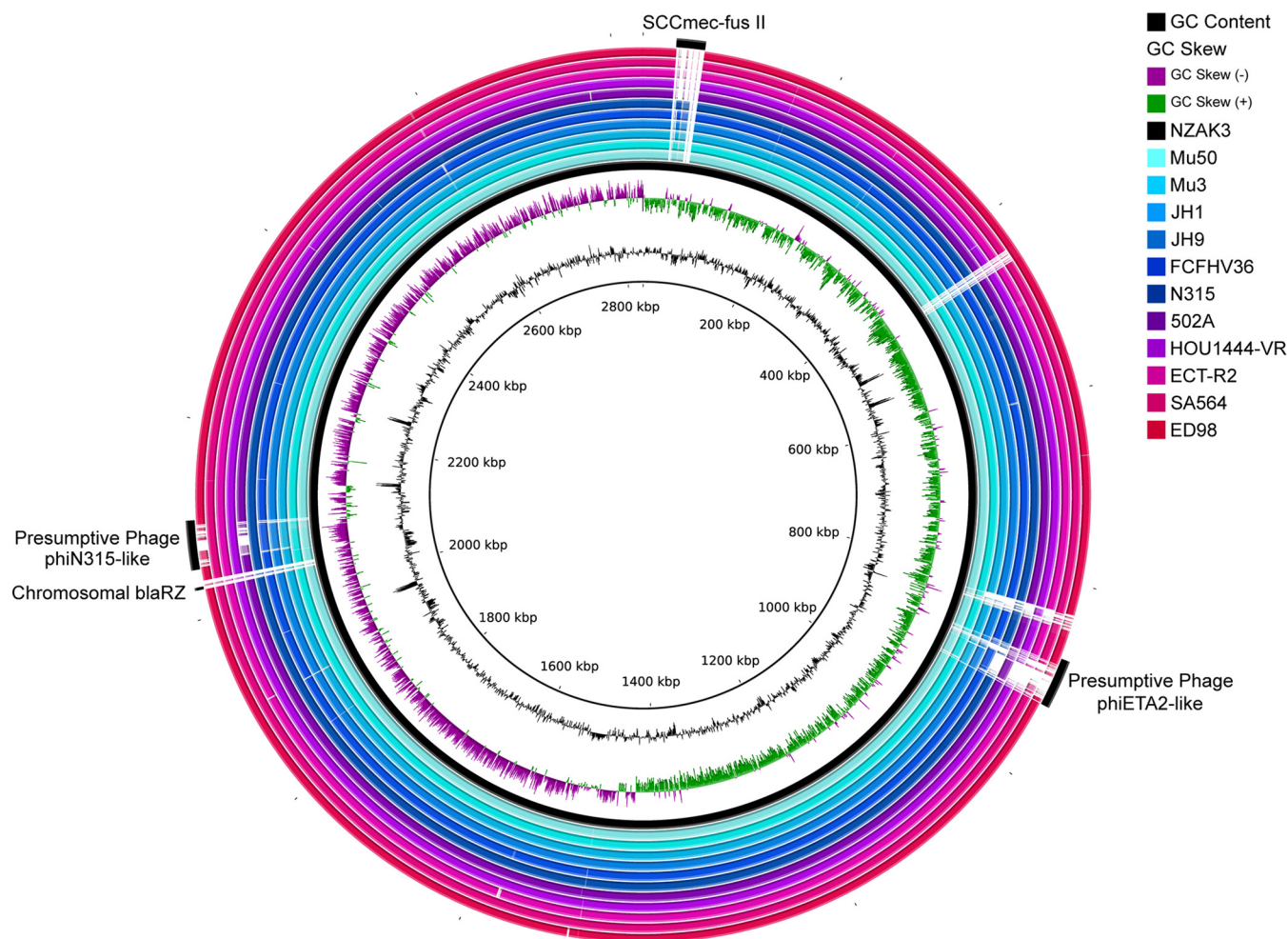


FIG 1 Sequence comparison of all complete genomes of clonal complex 5 (CC5) *S. aureus* strains. A BLAST-based comparison of all complete CC5 *S. aureus* genomes against representative FA-resistant strain NZAK3, visualized with BRIG, illustrates a high level of sequence conservation among members of the CC5 lineage. Genomic regions in NZAK3 that were variably present in other CC5 *S. aureus* strains include the *SCCmec-fus II* mobile element, a chromosomal copy of the *blaRZ* locus, and two complete phage regions, which were identified using PHAST as a presumptive phiETA2 ([NC_008798](#)) and phiN315 ([NC_004740](#)).

lates of the same ST typically grouped together in ST-specific clades, the only exception being ST5, which represented multiple distinct clades and singletons. In general, isolates from NZ clustered together, and only one isolate, recovered outside NZ, was found to cluster with an NZ-specific clade ([Fig. 2](#)). This NZ clade represented the AK3 ST5 MRSA clone, and all FA-resistant CC5 isolates recovered from NZ resided within this clade. The median core genome pairwise SNP distance between two isolates in this clade was only 49 SNPs, with a minimum distance of five SNPs and a maximum distance of 96 SNPs, highlighting the extremely restricted genetic diversity of FA-resistant ST5 MRSA in NZ. The single international isolate to cluster with the AK3 ST5 MRSA clade was an FA-resistant MRSA isolate (MRSA10) recovered in the UK during 2011 to 2012 ([6](#)). MRSA10 had a median core genome pairwise SNP distance of only 59 SNPs to other AK3 isolates. Interestingly, the other CC5 FA-resistant isolates from the United Kingdom formed a single ST149 clade ($n = 6$) and one ST5 singleton. Taken together, these findings suggest that AK3 ST5 MRSA forms a genetically distinct clone within ST5, as indicated by the genetic distance between the AK3 clade and representative

international ST5 isolates. Furthermore, the location of UK isolate MRSA10 in the AK3 clade suggests potential recent intercontinental transmission between these two countries.

The inclusion of FA-susceptible isolates in this population model also enabled an investigation into the origin of the AK3 ST5 MRSA clone. The phylogenetic proximity of the AK3 ST5 MRSA clade to four ST5 MSSA isolates recovered from NZ suggests that this clone emerged from locally circulating MSSA rather than being imported from overseas ([Fig. 2](#)), as has been described for ST5 MRSA in other settings ([23](#)). Furthermore, locally circulating ST149 *S. aureus* strains (belonging to CC5) also were found to be more similar at a core genome level to AK3 ST5 MRSA than to representative international lineages of ST5. The hypothesis that AK3 ST5 MRSA emerged from local ST5 MSSA also was supported by the absence of any remnants of an *SCCmec* element in the four NZ ST5 MSSA strains. As described below, the *fusC* gene was present in all FA-resistant AK3 ST5 MRSA isolates as part of a chimeric *SCCmec*-*SCC* mobile element integrated into the chromosome at *orfX*. The inspection of the region surrounding *orfX* found it to be intact in all four ST5 MSSA isolates (see Fig. S1 in the

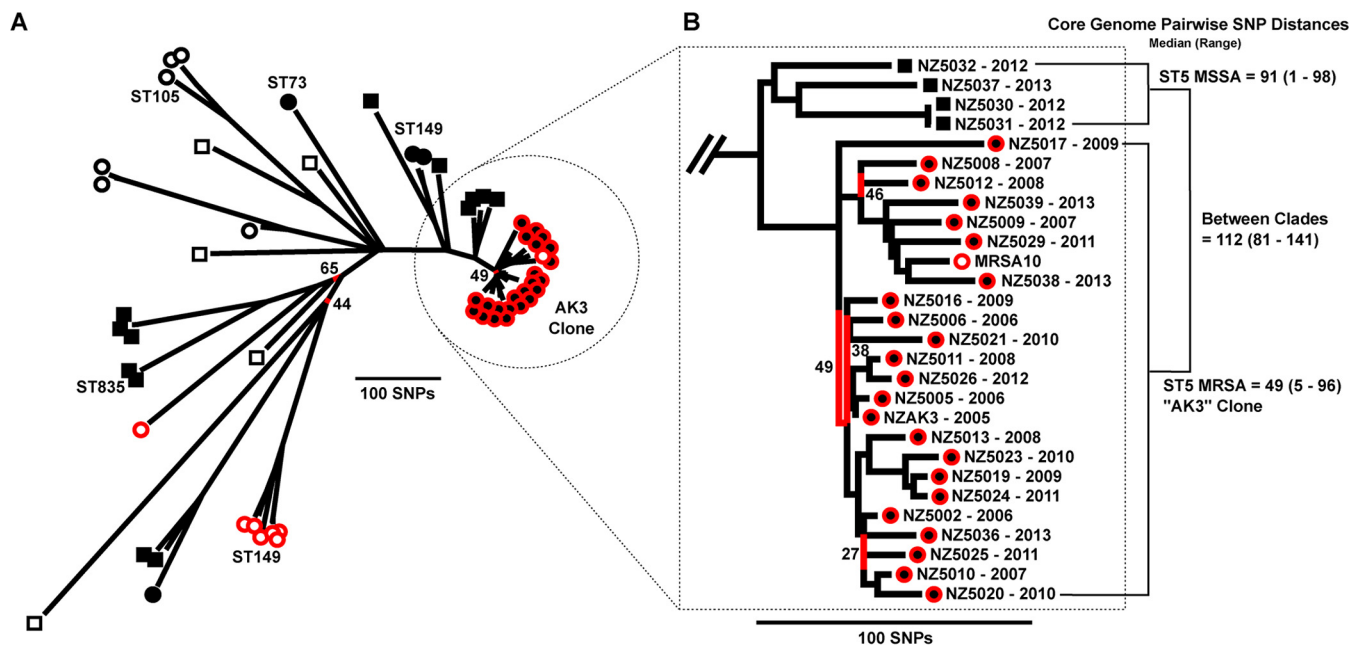


FIG 2 Population structure of fusidic acid-resistant CC5 *S. aureus*. (A) The maximum likelihood tree, constructed from core genome SNPs, illustrates the population structure of a representative global population of CC5 *S. aureus* strains. Terminal branch symbols indicate the location from which isolates were recovered and both methicillin and fusidic acid resistance phenotypes. Symbol shape: circle, MRSA; square, MSSA. Inner symbol color: black, recovered in New Zealand; white, recovered outside New Zealand. Outer symbol color: red, fusidic acid resistant; black, fusidic acid susceptible. (B) Enlarged view of the population structure for the AK3 ST5 MRSA clade and a locally circulating ST5 MSSA clade. The core genome pairwise SNP distances, calculated from the SNP matrix used to construct this tree, indicates the median and range of SNPs that separate any two isolates within the AK3 ST5 MRSA clade, the ST5 MSSA clade, and between the two clades. In both panels A and B, branches with less than 70% bootstrap support are colored red, and the exact support percentage is provided.

supplemental material), suggesting that these isolates had not previously carried an SCC element in their ancestral genome. Therefore, it is possible that the acquisition of *fusC* in the AK3 ST5 MRSA clone occurred more recently, following its divergence from locally circulating ST5 MSSA, although this hypothesis should be explored further in the framework of a larger sample of international ST5 MSSA.

Genetic context of *fusC* in ST1 and ST5 *S. aureus* isolates. In order to elucidate the genetic context of the *fusC* gene among NZ lineages of *S. aureus* and to compare it to other described CC5 and ST1 *fusC*-containing SCC elements, the genomes of *fusC*-harboring isolates were *de novo* assembled. These included 23 isolates belonging to the AK3 ST5 MRSA clade and six *fusC*-harboring ST1 *S. aureus* isolates recovered in New Zealand. In addition, six representative CC5 and ST1 *S. aureus* isolates from published studies were included. These isolates contained previously described *fusC*-containing mobile elements, including SCC₄₇₆ in MSSA476 (ST1) (4), and previously described elements from a recent study (6), namely, the composite elements SCCmecIVa-SCC₄₇₆ in isolate MRSA2 (ST1 MRSA), SCC₁₄₉-SCCmecIVa in isolate MRSA17 (ST149 MRSA), and the chimeric elements SCCmec-fusI in isolate MRSA3 (ST1 MRSA), SCCmec-fusII in isolate MRSA10 (ST5 MRSA), and SCCmec-fusIII in isolate MRSA7 (ST1 MRSA) (see Table S1 in the supplemental material). In a previous study by Ellington et al. (6), the chimeric elements SCCmec-fusII (29.1 kb) and SCCmec-fusIII (25.4 kb) were reported to have been identified in isolates MRSA7 (ENA accession number [ERR108061](#)) and MRSA10 (ENA accession number [ERR108040](#)), respectively. However, we found this to be the reverse, with the sequence reads available for isolate MRSA7 containing the shorter SCCmec-fusIII

element and for MRSA10 containing the longer SCC*mec*-*fusII* element. Subsequently, we refer to these chimeric elements with the annotations previously reported (6) but with the isolate names reversed.

Among ST5 MRSA isolates, reference strain NZAK3 contained a chimeric SCC*mec*IV-SCC₄₇₆ element spanning 29,120 bp that displayed 99.9% nucleotide sequence to the SCC*mec*-*fusII* element identified in isolate MRSA10 from the previous UK study (6) (Fig. 3). This finding further supports the likely intercontinental spread of this clone, as suggested by the population model. The examination of the draft assemblies of the other 23 AK3 ST5 MRSA isolates found that they all carried *fusC* as part of a chimeric SCC*mec*-SCC₄₇₆ element, although in 15 isolates this element was fragmented and located across two contigs. The longest fragment ranged in length from 78% to 87% of the full length of the *fusC*-containing chimeric SCC*mec*IV-SCC₄₇₆ and displayed >99.7% nucleotide sequence identity to the chimeric SCC*mec*IV-SCC₄₇₆ element found in NZAK3. Of the eight isolates with a complete SCC*mec*-SCC₄₇₆ element, seven displayed >99.9% nucleotide sequence identity to the chimeric SCC*mec*IV-SCC₄₇₆ element in NZAK3. The single AK3 ST5 MRSA isolate with a different chimeric SCC*mec*IV-SCC₄₇₆ element was isolate NZ5029, which displayed only 89.9% nucleotide identity to the chimeric SCC*mec*IV-SCC₄₇₆ element in NZAK3. A visual assessment of this alignment identified that the chimeric SCC*mec*IV-SCC₄₇₆ element in NZ5029 contained a previously unreported transposon insertion within the SCC*mec*IV portion upstream of *mecA* (Fig. 3). This transposon was flanked by two copies of IS431, and the intervening region encoded two genes annotated as a *repL* (firmicute plasmid replication protein) and *ksgA* (dimethyladenosine trans-

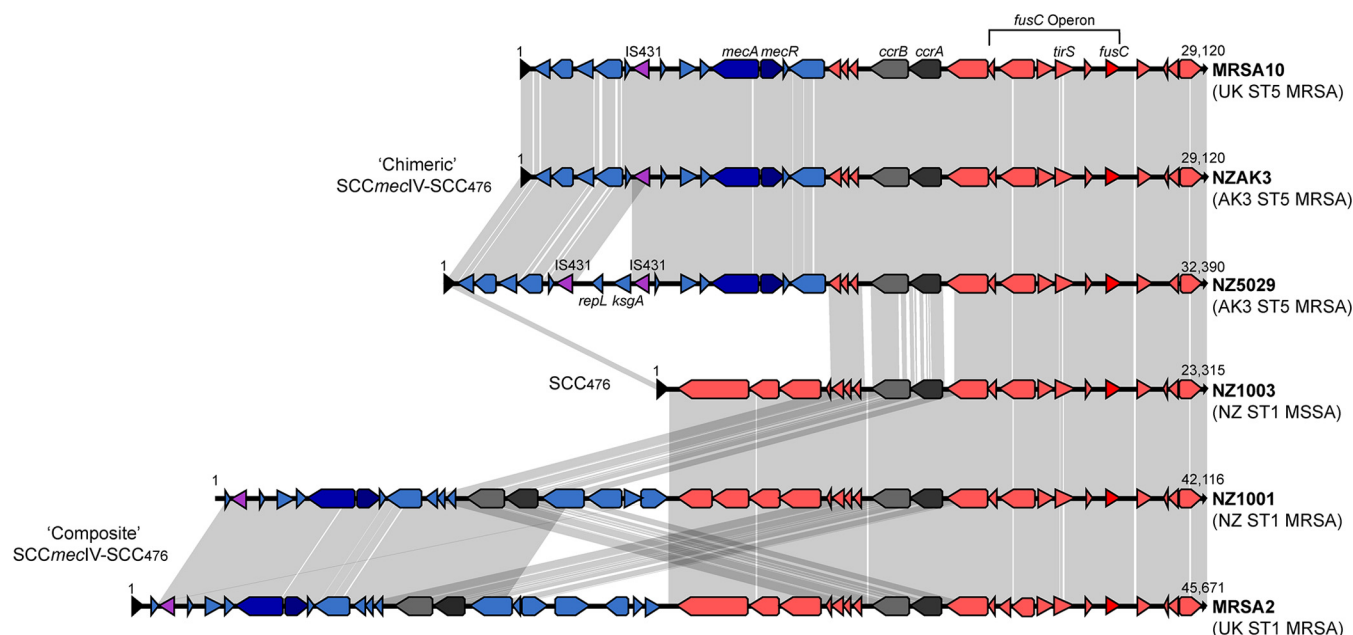


FIG 3 Comparison of *fusC*-containing staphylococcal cassette chromosome elements in ST1 and ST5 lineages of *S. aureus*. Schematic diagram illustrating the genetic organization of the SCC 476 (SCC₄₇₆) element in ST1 MSSA (represented by strain NZ1003), the chimeric SCCmecIV-SCC₄₇₆ element in ST5 MRSA (strain NZAK3), and the composite SCCmec-SCC₄₇₆ element in ST1 MRSA (strain NZ1001). The transposon-containing region within the chimeric SCCmecIV-SCC₄₇₆ element in strain NZ5029 also is illustrated. The SCCmec elements from MRSA2 and MRSA10 from a previous study (6) are included for comparison. Blue or purple arrows indicate sequences present within SCCmec regions. The direction of the arrows indicates the direction of transcription for open reading frames. Only coding sequences of >200 bp are shown. Gray shaded areas represent regions that share >99% nucleotide sequence identity.

ferase). This insertion was exclusive to this strain, as a nucleotide BLAST search for the *ksgA* gene returned no hits among the other AK3 ST5 MRSA isolates.

Importantly, the high level of nucleotide identity between *fusC*-containing SCC elements in AK3 ST5 MRSA isolates supports the hypothesis that the AK3 ST5 MRSA clone emerged following a single acquisition of the chimeric *fusC*-containing SCCmecIV-SCC₄₇₆ element with subsequent clonal expansion. Furthermore, the colocation of *fusC* in the same mobile element as *mecA* is of particular concern in NZ, where the extensive use of topical FA in NZ appears not only to be selecting for FA-resistant *S. aureus* but also coselecting for MRSA. Therefore, the genotypic association of *fusC* with *mecA* has important implications for the emergence of MRSA clones in populations with high usage of fusidic acid.

In addition to ST5 MRSA, the two other dominant FA-resistant *S. aureus* lineages identified in NZ were ST1 MSSA and ST1 MRSA (2). To determine the genetic context of the *fusC* gene in these clones, six representative ST1 isolates were sequenced. Of these six isolates, five were MSSA and all carried *fusC* as part of an SCC element that displayed >99.8% nucleotide sequence identity to the previously described 23.3-kb SCC₄₇₆ element identified in strain MSSA476 (4). In a recent survey, FA-resistant ST1 MSSA was the dominant MSSA clone circulating in the NZ community, accounting for 15% of all MSSA strains (2). This further supports the notion that high population use of FA has selected for dominant FA-resistant clones, resulting in marked changes in the molecular epidemiology of *S. aureus* in NZ (2, 3). The single ST1 MRSA isolate contained a composite SCCmec-SCC₄₇₆ element (42,116 bp) but had only 88.9% nucleotide sequence identity to the only previously described composite element in ST1 MRSA,

the SCCmecIVa-SCC₄₇₆ element previously identified in isolate MRSA2 (6) (Fig. 3).

A comparison of the four different SCC elements identified in ST5 and ST1 *S. aureus* isolates in NZ and the two genetically similar SCC elements identified in isolates from the United Kingdom is presented in Fig. 3. The *fusC* gene was located consistently within the joining (J1) region between the *ccrAB2* genes and the downstream direct-repeat (DR) region of an SCC₄₇₆ type element. Of note, among these six SCC elements, there was an ~5-kb region with >99% nucleotide sequence identity upstream of *fusC*, suggesting a common ancestral origin for this specific sequence. Moreover, within this region was a *tirS* gene encoding a TIR domain named TirS. Recent work has suggested that TirS contributes to bacterial virulence by interfering with TLR2 signaling pathways and enhancing bacterial survival within the host (24). It is possible that, in addition to the presence of the *fusC* gene, the presence of TirS also confers a selective advantage for these dominant clones, although it is presently unknown whether clinical infections caused by TirS-producing *S. aureus* strains are more severe than those of TirS-deficient isolates. However, it is of concern that, in addition to the presence of genes encoding antimicrobial resistance, there is also evidence of horizontal transfer of a collocated putative virulence factor.

Association of *fusC* with SCC-type elements in diverse lineages of *S. aureus*. In this study, we found that *fusC* was invariably present within an SCC element, with or without the *mecA* gene. This also was found to be the case for other published *fusC*-harboring sequences deposited in the GenBank database, including SCC elements from Ireland (5) and Taiwan (7). This finding suggests that, similar to *mecA* transmission, SCC-mediated horizontal gene transfer is the major mechanism of *fusC* dissemination.

This is distinct from the mechanism of dissemination of the other major acquired FA resistance gene, *fusB*, which is transferred primarily on plasmids, such as pUB101 and pUB102 (25, 26) and p11819-97, among CC80 *S. aureus* strains in Europe (27).

Of further interest was our finding that, in all previously identified *fusC*-containing SCC elements, both in this study and in other published sequences (4–7), *fusC* was always present within a conserved ~1.3-kb region, which shared one upstream putative CDS (Fig. 2). A BLASTX and Pfam search did not assign any functional roles for these putative proteins. However, given the consistent finding of this CDS in association with *fusC*, it is possible that this protein plays a role in the regulation or modulation of *fusC* transcription. Future work should attempt to elucidate the function of these proteins in relation to *fusC* expression.

Conclusions. In this study, we described the emergence and possible origins of fusidic acid-resistant AK3 ST5 MRSA, the most prevalent MRSA clone in NZ. On the basis of population-based comparative genomics, it is likely that this clone has emerged from locally circulating ST5 MSSA strains in New Zealand. The AK3 ST5 MRSA clone carries a chimeric SCCmecIV-SCC₄₇₆ element harboring *fusC*, which appears to have been acquired once and maintained among members of this clone. The invariable presence of *fusC* within an SCC-type element indicates that SCC-mediated horizontal transfer is the primary mechanism for dissemination of *fusC*. The genotypic association of *fusC* with *mecA* has important implications for the emergence of MRSA in populations with high usage of fusidic acid, and the colocation of *fusC* with *tirS* may confer an additional fitness advantage. The finding of AK3 ST5 MRSA in the United Kingdom suggests intercontinental transmission, which should be of concern to other regions that use topical FA to treat superficial SSTI. This study has revealed the rapidity and process by which an *S. aureus* population evolved and expanded in response to a human therapeutic intervention, in this case the increased use of topical FA. In the face of these data and those of related studies, it is clear that all of our efforts to control bacterial infections with antibiotics, disinfectants, and even vaccines are being met by swift Darwinian responses. The use of genomics for the prospective surveillance of clinical bacterial populations will be a critical monitoring tool in an imminent postantibiotic era.

ACKNOWLEDGMENT

We acknowledge the technical assistance of the staff of the Antimicrobial Reference Laboratory at the Institute of Environmental Science and Research.

REFERENCES

1. Stevens DL, Bisno AL, Chambers HF, Everett ED, Dellinger P, Goldstein EJ, Gorbach SL, Hirschmann JV, Kaplan EL, Montoya JG, Wade JC. 2005. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. *Clin Infect Dis* 41:1373–1406. <http://dx.doi.org/10.1086/497143>.
2. Williamson DA, Monecke S, Heffernan H, Ritchie SR, Roberts SA, Upton A, Thomas MG, Fraser JD. 2014. High usage of topical fusidic acid and rapid clonal expansion of fusidic acid-resistant *Staphylococcus aureus*: a cautionary tale. *Clin Infect Dis* 59:1451–1454. <http://dx.doi.org/10.1093/cid/ciu658>.
3. Williamson DA, Roberts SA, Ritchie SR, Coombs GW, Fraser JD, Heffernan H. 2013. Clinical and molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in New Zealand: rapid emergence of sequence type 5 (ST5)-SCCmec-IV as the dominant community-associated MRSA clone. *PLoS One* 8:e62020. <http://dx.doi.org/10.1371/journal.pone.0062020>.
4. Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci U S A* 101:9786–9791. <http://dx.doi.org/10.1073/pnas.0402521101>.
5. Kinnevey PM, Shore AC, Brennan GI, Sullivan DJ, Ehrlich R, Monecke S, Slickers P, Coleman DC. 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome mec (SCCmec)-SCC-SCCRISPR composite element in Irish hospitals. *Antimicrob Agents Chemother* 57:524–531. <http://dx.doi.org/10.1128/AAC.01689-12>.
6. Ellington MJ, Reuter S, Harris SR, Holden MT, Cartwright EJ, Greaves D, Gerver SM, Hope R, Brown NM, Török ME, Parkhill J, Köser CU, Peacock SJ. 2015. Emergent and evolving antimicrobial resistance cassettes in community-associated fusidic acid and methicillin-resistant *Staphylococcus aureus*. *Int J Antimicrob Agents* 45:477–484. <http://dx.doi.org/10.1016/j.ijantimicag.2015.01.009>.
7. Lin YT, Tsai JC, Chen HJ, Hung WC, Hsueh PR, Teng LJ. 2014. A novel staphylococcal cassette chromosomal element, SCC_{fusC}, carrying *fusC* and *speG* in fusidic acid-resistant methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 58:1224–1227. <http://dx.doi.org/10.1128/AAC.01772-13>.
8. Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* 357:1225–1240. [http://dx.doi.org/10.1016/S0140-6736\(00\)04403-2](http://dx.doi.org/10.1016/S0140-6736(00)04403-2).
9. Parker D, Narechania A, Sebra R, Deikus G, Larussa S, Ryan C, Smith H, Prince A, Mathema B, Ratner AJ, Kreiswirth B, Planet PJ. 2014. Genome sequence of bacterial interference strain *Staphylococcus aureus* 502A. *Genome Announc* 2:e00284–14.
10. Lindqvist M, Isaksson B, Grub C, Jonassen TO, Hallgren A. 2012. Detection and characterisation of SCCmec remnants in multiresistant methicillin-susceptible *Staphylococcus aureus* causing a clonal outbreak in a Swedish county. *Eur J Clin Microbiol Infect Dis* 31:141–147. <http://dx.doi.org/10.1007/s10096-011-1286-y>.
11. Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, Cartwright RA, Simpson AJ, Rambaut A, Nubel U, Fitzgerald JR. 2009. Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 106:19545–19550. <http://dx.doi.org/10.1073/pnas.0909285106>.
12. Panesso D, Planet PJ, Diaz L, Hugonnet JE, Tran TT, Narechania A, Munita JM, Rincon S, Carvajal LP, Reyes J, Londono A, Smith H, Sebra R, Deikus G, Weinstock GM, Murray BE, Rossi F, Arthur M, Arias CA. 2015. Methicillin-susceptible, vancomycin-resistant *Staphylococcus aureus*, Brazil. *Emerg Infect Dis* 21:1844–1848. <http://dx.doi.org/10.3201/eid2110.141914>.
13. Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, Bruce D, Rubin E, Myers E, Siggia ED, Tomasz A. 2007. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc Natl Acad Sci U S A* 104:9451–9456. <http://dx.doi.org/10.1073/pnas.0609839104>.
14. Neoh HM, Cui L, Yuzawa H, Takeuchi F, Matsuo M, Hiramatsu K. 2008. Mutated response regulator graR is responsible for phenotypic conversion of *Staphylococcus aureus* from heterogeneous vancomycin-intermediate resistance to vancomycin-intermediate resistance. *Antimicrob Agents Chemother* 52:45–53. <http://dx.doi.org/10.1128/AAC.00534-07>.
15. Giraud C, Hausmann S, Lemeille S, Prados J, Redder P, Linder P. 2015. The C-terminal region of the RNA helicase CshA is required for the interaction with the degradosome and turnover of bulk RNA in the opportunistic pathogen *Staphylococcus aureus*. *RNA Biol* 12:658–674. <http://dx.doi.org/10.1080/15476286.2015.1035505>.

16. McCulloch JA, Silveira AC, Lima Moraes Ada C, Perez-Chaparro PJ, Ferreira Silva M, Almeida LM, d'Azevedo PA, Mamizuka EM. 2015. Complete genome sequence of *Staphylococcus aureus* FCFHV36, a methicillin-resistant strain heterogeneously resistant to vancomycin. *Genome Announc* 3:e00893–15.
17. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <http://dx.doi.org/10.1101/gr.2289704>.
18. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.
19. Guindon S, Delsuc F, Dufayard JF, Gascuel O. 2009. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* 537:113–137. http://dx.doi.org/10.1007/978-1-59745-251-9_6.
20. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948. <http://dx.doi.org/10.1093/bioinformatics/btm404>.
22. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. *Bioinformatics* 21:3422–3423. <http://dx.doi.org/10.1093/bioinformatics/bti553>.
23. Nubel U, Roumagnac P, Feldkamp M, Song JH, Ko KS, Huang YC, Coombs G, Ip M, Westh H, Skov R, Struelens MJ, Goering RV, Strommenger B, Weller A, Witte W, Achtman M. 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 105:14130–14135. <http://dx.doi.org/10.1073/pnas.0804178105>.
24. Askarian F, van Sorge NM, Sangvik M, Beasley FC, Henriksen JR, Sollid JU, van Strijp JA, Nizet V, Johannessen M. 2014. A *Staphylococcus aureus* TIR domain protein virulence factor blocks TLR2-mediated NF-kappaB signaling. *J Innate Immun* 6:485–498.
25. Monecke S, Slickers P, Ehrlich R. 2008. Assignment of *Staphylococcus aureus* isolates to clonal complexes based on microarray analysis and pattern recognition. *FEMS Immunol Med Microbiol* 53:237–251. <http://dx.doi.org/10.1111/j.1574-695X.2008.00426.x>.
26. O'Brien FG, Price C, Grubb WB, Gustafson JE. 2002. Genetic characterization of the fusidic acid and cadmium resistance determinants of *Staphylococcus aureus* plasmid pUB101. *J Antimicrob Chemother* 50:313–321. <http://dx.doi.org/10.1093/jac/dkf153>.
27. Stegger M, Wirth T, Andersen PS, Skov RL, De Grassi A, Simoes PM, Tristan A, Petersen A, Aziz M, Kiil K, Cirkovic I, Udo EE, del Campo R, Vuopio-Varkila J, Ahmad N, Tokajian S, Peters G, Schaumburg F, Olsson-Liljequist B, Givskov M, Driebe EE, Vigh HE, Shittu A, Ramdani-Bougessa N, Rasigade JP, Price LB, Vandenesch F, Larsen AR, Laurent F. 2014. Origin and evolution of European community-acquired methicillin-resistant *Staphylococcus aureus*. *mBio* 5:e01044–14. <http://dx.doi.org/10.1128/mBio.01044-14>.

3.2.2 Supplementary Material

Baines, S. L., B. P. Howden, H. Heffernan, T. P. Stinear, G. P. Carter, T. Seemann, J. C. Kwong, S. R. Ritchie and D. A. Williamson (2016). "Rapid emergence and evolution of *Staphylococcus aureus* clones harboring *fusC*-containing staphylococcal cassette chromosome elements." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.03020-15.

Available in Appendix 1: Supplementary Dataset S1. Demographic and genome sequencing information for all bacterial isolates used in this study.

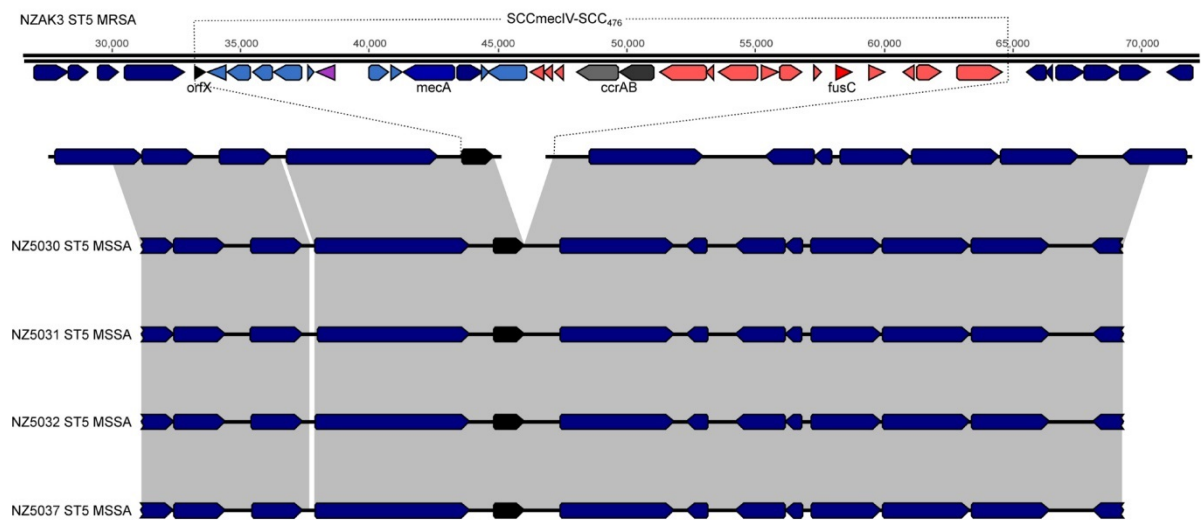


FIG S1 Alignment of region surrounding *orfX* demonstrating the intact *orfX* gene in ST5 MSSA from New Zealand that are closely related to NZAK3.

3.3 Extended Analysis

The original publication described the evolution of the AK3 MRSA clone in New Zealand, a FA-resistant lineage that has recently emerged in the region and become the dominant MRSA clone in both community and healthcare environments. One of the limitations of this study was the inclusion of a limited number of sequences representing international isolates, reducing the phylogeographic inferences that could be made. As with the previous chapter, sequence data representing 376 CC5 *S. aureus* had become available since this work was published. This extended analysis was conducted to revisit the evolution of AK3 MRSA in a more globally diverse dataset and continue to explore the association of *fusC* with the SCC element.

3.3.1 Materials & Methods

Bacterial Isolates. This extended analysis utilised the genome sequence data for a total of 408 CC5 *S. aureus* isolates. The source of these isolates and selection criteria are detailed below, and isolate specific information is available in Appendix 2. For all unpublished isolates detail is provided about the storage, genomic DNA extraction and sequencing conditions.

Published sequence data

- The complete genomes of 21 CC5 *S. aureus* isolates were sourced from (40, 45, 194, 348-356). The complete genome of MSSA476 (ST1 MSSA) was also included (47).
- Thirty-nine isolates were sourced from the publication provided in section 3.2 (194). All CC5 isolates from this study were included.
- One hundred and fifty-two isolates were sourced from the collections of (193, 331-334, 357-360). All CC5 isolates that could be accessed from these studies were included.
- Thirty-five isolates were sourced from the collection of (361). A limited number of CC5 isolates were included and selected to maximise the temporal and geographic diversity of this collection.

Unpublished sequence data

- The sequence data for 22 isolates were provided by the Institute of Environmental Science and Research, New Zealand. These represented additional CC5 *S. aureus* recovered from clinical cases of SSTI in New Zealand around the same time as the original collection (2006 to 2013).
- Sixty-three isolates were sourced from the ANZCOSS study (335). All CC5 isolates from unique patients selected for an ANZCOSS sub-study were included (306).
- Forty-five isolates were sourced from the 2015 ASSOP collection. Sequence data for all CC5 isolates was kindly provided by the Australian Group on Antimicrobial Resistance (<http://agargroup.org.au/agar-surveys>).
- Thirty isolates were sourced from the VANESSA study (336). All CC5 isolates from unique patients were included.

Storage & Whole Genome Sequencing. Bacterial isolates sourced from the ANZCOSS and VANESSA collections were stored and sequenced as described in section 2.3.1.

Genome Assembly and Detection of *fusC*. *De novo* assembly was performed for all 386 isolates with short-read sequence data using SPAdes v3.11.0 (Bankevich, Nurk et al. 2012), excluding the pre-assembly error correction (“--only-assembler”), including post-assembly error correction (“--careful”) and a minimum coverage threshold of five (“--cov-cutoff”) to remove potential contaminants. The detection of *fusC* and recovery of the surrounding genomic context was performed using contig-puller (<http://github.com/kwongj/contig-puller/>). Contigs that contained a match to the *fusC* sequence of reference NZAK3 (Genbank accession LT009690), with a nucleotide identity threshold of $\geq 98\%$, were aligned and annotated to create a multi-genbank file that was investigated using Geneious v8.1.5 (Biomatters).

Phylogenetic Analysis. The Illumina short-read data for all 407 CC5 *S. aureus* isolates was mapped to reference genome NZAK3, using Snippy v3.2-dev (<https://github.com/tseemann/snippy>). In the case of the 20 complete genomes, pseudo short-reads (150 bp paired-end) were generated prior to mapping using the option “--ctgs” in Snippy v3.2-dev. A whole genome alignment was generated using a minimum coverage of 10 reads and base-call stringency of 90% to call nucleotide variants. Examination of this alignment found one isolate to be highly divergent (data not shown). This isolate (“87”) was excluded from further analysis. The whole genome alignment was reconstructed with the remaining 406 isolates. Recombination was assessed using ClonalFrameML v.1.11 (338). However as so few recombinant events were detected, masking these sites prior to phylogenetic analysis was not required. Testing of nucleotide substitution models was performed on the whole genome alignment using IQ-TREE v1.6.1 (339), with the GTR nucleotide substitution model producing the highest likelihood. A second alignment was generated by mapping the short-read data for all 50 *fusC*-harbouring isolates against the SCC₄₇₆ sequence of *S. aureus* MSSA476 (Genbank accession BX571857). ML phylogenetic trees were constructed from both alignments using IQ-TREE v1.6.1, run under a “GTR+G4” model with 10,000 bootstrap replicates. The phylogeny of all isolates was rooted using information gained by the inclusion of four diverse *S. aureus* genomes (data not shown). A mid-point root was applied to the SCC₄₇₆ phylogeny. All alterations to the aesthetics of the phylogenetic trees presented were performed in either FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>) or Inkscape v0.91 (<https://inkscape.org/en/>).

3.3.2 Results & Discussion

To re-examine the origins of the AK3 MRSA clones in New Zealand, the CC5 collection examined in (194) was supplemented with the sequence data of an additional 376 local and international CC5 isolates. This collection of 406 isolates represented a time span of 33 years (1982 to 2015) and 15 countries across six continents: Australasia (n = 200, Australia and New Zealand), North America (n = 91, United States), Europe (n = 90, United Kingdom and Ireland, Germany, Sweden, and Switzerland), Africa (n = 14, Gambia and Ghana), Asia (n = 9, Japan and Thailand), South America (n = 1, Brazil), and one isolate of unknown geographic origin.

Clonal Complex 5 represents a large genetically diverse and geographically dispersed lineage of *S. aureus* with highly successful hospital- and community-associated clones (69, 141, 362). A maximum likelihood phylogenetic tree, inferred from the whole genome alignment of 406 CC5 isolates, is presented in Figure 3.1. It has been previously proposed that the evolution of CC5 MRSA is dominated by the frequent loss and acquisition of the *SCCmec* element into geographically established and genetically divergent CC5 backgrounds, which appear as distinct geographic clades encompassing both MRSA and MSSA isolates in a phylogenetic model (362). This is in contrast to an alternative hypothesis in which regular intercontinental transmission seeds CC5 MRSA into new regions. In this instance, phylogenetic clades would be observed to have a stronger association with the presence of *SCCmec* over geographic origin. The population structure illustrated in Figure 3.1 showed a mixture of these models, with some clades demonstrating a strong geographic signal and mixed MRSA/MSSA genotypes, while others demonstrate one predominant genotype with a mixture of geographic sources. This finding is both consistent with the publication by (362), but suggests that intercontinental transmission may have a greater impact on the evolution of CC5 than previously thought. However, because this collection included isolates from a range of studies with different and potentially biased sampling approaches, further work is needed to comprehend the true extent to which intercontinental transmission may influence the global dissemination of CC5 *S. aureus*.

In this population model, the CC5 isolates from New Zealand largely fell within a single clade (Figure 3.1A), which included all 37 *fusC*-harbouring isolates. All New Zealand isolates that clustered elsewhere in the tree did not harbour *fusC*. Three additional clusters of *fusC*-harbouring isolates were identified. This included a ST149 clade, predominantly composed of isolates from the United Kingdom, that had been previously described in the work of (193), and had also been observed in the original publication (194). The other two groups of *fusC*-harbouring isolates were all ST5 MRSA, and were recovered from Australia, Germany and the United Kingdom. No *fusC*-harbouring CC5 MSSA were identified in this collection.

Consistent with the original publication, closer inspection of the population structure surrounding the 50 *fusC*-harbouring isolates revealed that all New Zealand isolates formed a single monophyletic clade, representing the AK3 clone (Figure 3.2A). This clade was most closely related to a clade of ST5 MSSA from New Zealand. All AK3 isolates contained nearly identical copies of the chimeric *SCCmecIVa-fusC* II element (Figure 3.2B, Table 3.1), apart from the single isolate that was described in the original publication to carry a structural variant of *SCCmecIVa-fusC* II (194). These data further supported that the AK3 clone had emerged locally, following the acquisition of a *SCCmecIVa-fusC* II element into a local circulating ST5 MSSA lineage. The FA resistance mediated by the acquisition of *fusC* has likely provided AK3 MRSA with an adaptive advantage, allowing it to spread in a community with high FA use.

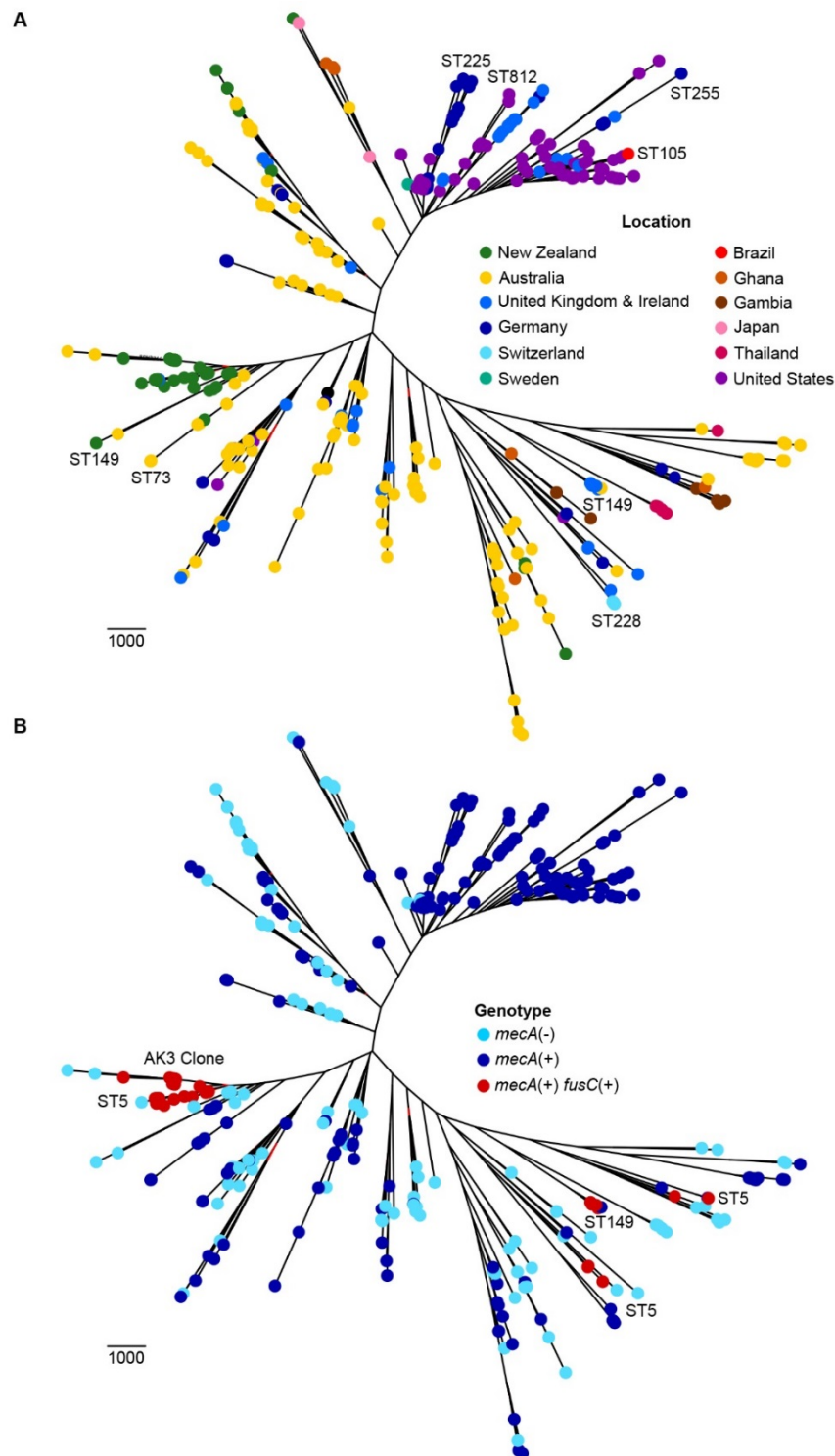


Figure 3.1 Global phylogeny of CC5 *S. aureus*. Maximum likelihood tree, inferred from the whole genome alignment of 406 isolates, illustrates the global population structure of CC5 *S. aureus*. Branches with < 70% bootstrap support are coloured red. Branch tips are coloured based on (A) geographic region of isolation, and (B) presence or absence of *mecA* and *fusC* genes (refer to key).

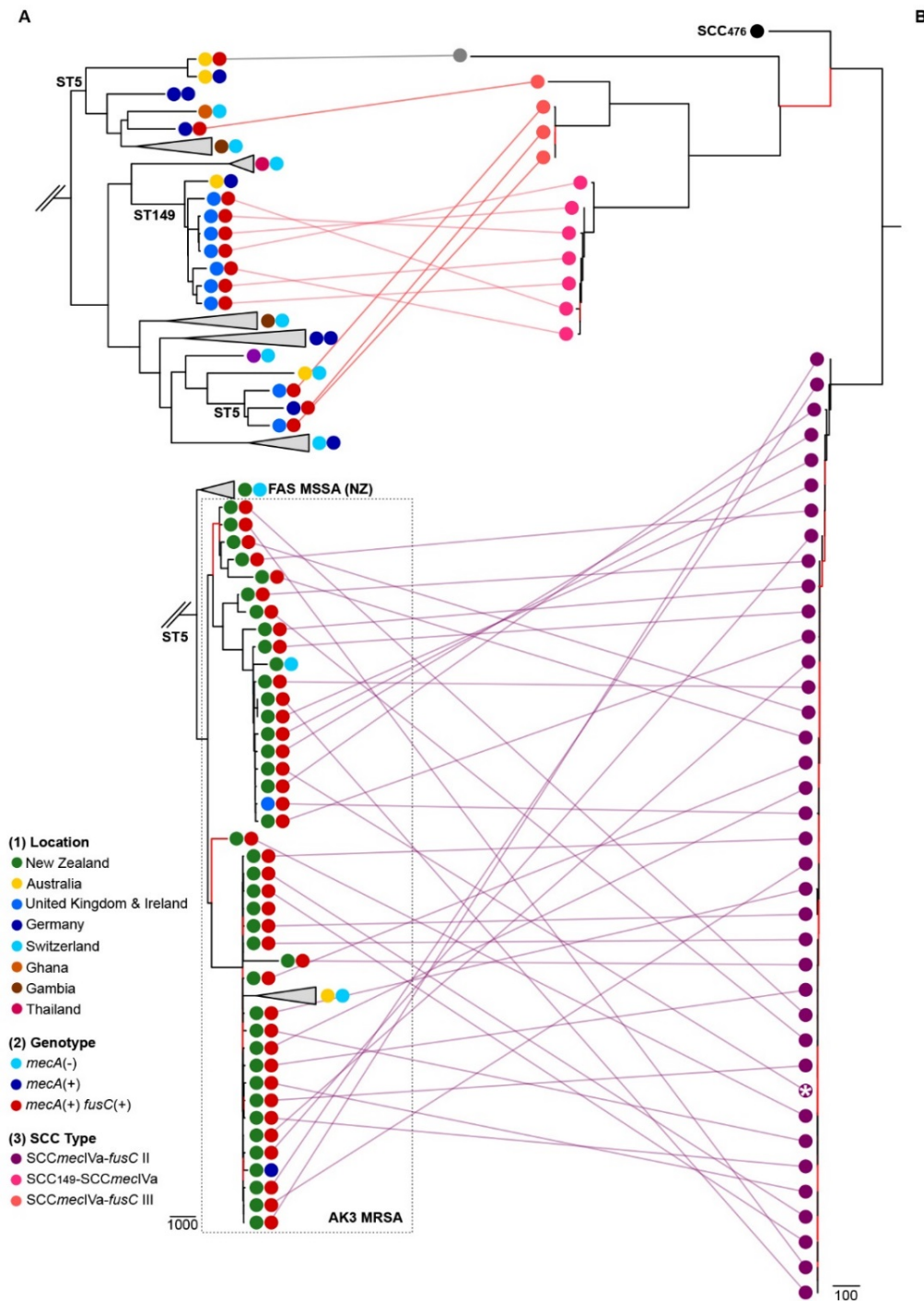


Figure 3.2 Comparison of the population structure of fusidic acid resistant CC5 MRSA and SCC₄₇₆. (A) The ML phylogeny is that from Figure 3.1, with only the regions containing *fusC*-harbouring isolates displayed. Branch tips are coloured to represent the geographic region of isolation (first circle), and the presence or absence of *mecA* and *fusC* genes (second circle, refer to key). Clades of non-*fusC* containing MRSA or MSSA with identical geographic source have been collapsed (grey wedge). (B) Presented is a ML phylogeny, inferred from the whole SCC₄₇₆ alignment of all 50 *fusC*-harbouring CC5 MRSA isolates. Branch tips are coloured to represent the type of SCC element (refer to key). In both phylogenetic trees, branches with < 70% bootstrap support are coloured red. Connections between isolates and their respective SCC element are illustrated. White star indicates the location of the structural variant of SCC*mecIVa-fusC* II.

Of interest in the original publication was the identification of the isolate MRSA10, a FA-resistant ST5 MRSA recovered in the United Kingdom between 2011 and 2012, that clustered with the AK3 clade (193, 194). This finding indicated that intercontinental transmission of the AK3 clone had occurred between New Zealand and the United Kingdom. Although no further isolates from the United Kingdom were found in the AK3 clade, two isolates from Australia were identified (Figure 3.2). Both were FA-susceptible MSSA and their location with the AK3 clade indicated that intercontinental transmission of this CC5 lineage had also occurred between New Zealand and Australia. These two isolates lacked the *SCCmecIVa-fusC* II element, accounting for the variant phenotype. There was no detectable disruption to the region surrounding *orfX* to suggest previous carriage of a *fusC*-containing SCC element. Although, SCC elements are capable of seamless excision, complicating attempts to predict previous carriage. Subsequently, the direction of transmission could not confidently be inferred in this case and it remains unclear whether these isolates represented intercontinental transmission of the AK3 clone with the subsequent loss of the *SCCmecIVa-fusC* II element or transmission of the ancestral FA-susceptible MSSA lineage from which AK3 had emerged. Temporal dating of the ancestral nodes for these isolates could not be performed due to the absence of a strong molecular clock signal in this dataset.

Examination of the other FA-resistant MRSA isolates indicated that they represented four unique SCC acquisition events (Figure 3.2B). All FA-resistant ST149 MRSA carried the composite *SCC₁₄₉-SCCmecIVa* element, previously describe by (193). As these isolates formed a monophyletic clade in both the whole genome phylogeny and the *SCC₄₇₆* phylogeny, it suggested a single acquisition of *SCC₁₄₉-SCCmecIVa* followed by the local expansion of this clone in the United Kingdom. A single ST149 isolate from Australia was also identified in this clade but it lacked the *SCC₁₄₉-SCCmecIVa* element. The remaining clade in the *SCC₄₇₆* phylogeny represented the *SCCmecIVa-fusC* III element, also previously describe by (193). This element appeared to have been acquired by genetically distinct ST5 isolates on at least two independent occasions (Figure 3.2). A single *fusC*-harbouring SCC element was identified in an isolate recovered from Australia (Figure 3.2). It was genetically distinct from the other *fusC*-harbouring SCC elements but could not be classified due to fragmentation in the assembly.

Consistent with previous studies, the *fusC* gene was invariably located within either a composite (*SCC₁₄₉-SCCmecIVa*) or chimeric (*SCCmecIVa-fusC* II and *SCCmecIVa-fusC* III) SCC element in all FA-resistant MRSA in this collection. This finding further supports that SCC-mediated HGT is likely the sole mechanism of dissemination for *fusC* in *S. aureus*. In CC5, *fusC* is consistently co-located with *mecA*, the genetic mediator of methicillin resistance, and *tirS*, a virulence associated gene which modulates the eukaryotic immune response (250). Co-location of these genes has important implications for the emergence of multidrug resistant *S. aureus* with potential enhanced immune evasion properties in communities with high FA use, and with what has been observed in New Zealand, can rapid alter the local epidemiology of *S. aureus*.

Table 3.1. Classification of *fusC*-harbouring CC5 *S. aureus* in this collection

	Isolates (n=50)	ST	Geographic Location (n)
Composite Elements			
SCC ₁₄₉ -SCCmecIVa	7	149	United Kingdom (7)
Chimeric Elements			
SCCmecIVa- <i>fusC</i> II	38	5	New Zealand (37), United Kingdom (1)
SCCmecIVa- <i>fusC</i> III	4	5	United Kingdom (2), Germany (2)
Other			
Unknown	1	5	Australia (1)

3.4 Conclusion

The emergence and spread of FA-resistant *S. aureus* poses a significant threat to the prevention and management of staphylococcal disease, especially SSTI. The work presented in this chapter provides insight into the evolutionary dynamics surrounding the emergence and spread of FA-resistant *S. aureus* in New Zealand, with a focus on ST5 MRSA. The recent and very rapid emergence of the AK3 clone appears to be the result of local acquisition of a *fusC*-harbouring SCCmec element (SCCmecIVa-*fusC* II) followed by extensive local expansion; the high-level and increasing use of FA in the region having fostered an environment in which AK3 has an adaptive advantage over FA-susceptible *S. aureus* clones. Although not reported in this chapter, both the FA-resistant ST1 MRSA and MSSA clones circulating in New Zealand exhibit similar evolutionary characteristics (228), illustrating the substantial influence unregulated FA use has had on the local *S. aureus* population. Intercontinental transmission may not have played a significant role in the spread of FA-resistant CC5 MRSA globally, but it has occurred and if more countries opt to licence and allow unrestricted use of FA it will likely contribute to the ongoing evolution of FA-resistant *S. aureus*. The convergent evolution of FA-resistant CC5 MRSA in New Zealand and the United Kingdom demonstrates the strong selective pressure imposed by the use of this topical antimicrobial. Furthermore, its use in these countries has led to the development of MGEs that co-locate *fusC*, *mecA*, *tirS*, and other antibiotic resistance encoding genes. While co-location of multiple resistance and virulence-associated genes is not a novel finding, nor is it an uncommon occurrence in pathogenic bacteria, it has significant consequences; the use of a single antibiotic can theoretically select for the combination of genes and promote the spread of lineages that are multidrug resistant and potential more pathogenic.

Chapter 4

**Remodelling of pSK1 family plasmids and enhanced
chlorhexidine tolerance in a dominant hospital lineage of
methicillin-resistant *Staphylococcus aureus***

4.1 Introduction

Plasmids play a central role in microbial evolution as they can mediate the transfer and dissemination of genes within and between populations, allowing for rapid adaptation to new or changing environments (363, 364). However, there are multiple challenges to studying plasmid sequences. (i) Plasmids are mobile resulting in their often-variable presence within a population. (ii) Plasmid architecture is significantly more flexible than the chromosome and is capable of drastic structural change in a short period of time. Further, (iii) the evolutionary history or the gene content of a plasmid may not reflect that of the bacterial isolate from which it was recovered, but rather the multiple clones and species that it has associated with (365-367).

In Australia, the two genetically distinct ST239 MRSA clades, the Australian and Asian-Australian clades, are known to harbour plasmids, specifically multidrug resistant plasmids that contribute to the antimicrobial resistance profiles of the two populations. The Australian clade was first recognised because it displayed resistance to gentamicin (296), a phenotypic feature that distinguished it from the MRSA clones circulating at the time, later identified as ST250 MRSA (368). Gentamicin resistance in these early ST239 isolates was encoded for by a bifunctional acetyltransferase-phosphotransferase [*aac(6')-aph(2'')*], commonly carried on a plasmid called pSK1 (215, 289, 369, 370). The pSK1 plasmid has been the focus of multiple published studies and is now recognised as the prototype for a family of intermediate sized (20-40 kb), theta-replicating, class II staphylococcal plasmids, and have been recovered in Australia and the United Kingdom (71, 289, 297, 371-373). Multiple antibiotic resistance and biocide tolerance mechanism are variably encoded for on pSK1 family plasmids, including an insensitive dihydrofolate reductase [*dhfrA*] (230, 280), a penicillinase [*blaZ*, carried as part of Tn552] (288, 372, 374), and a quaternary ammonium compound (QAC) multi-efflux pump [*qacA*] (198, 291).

The pSK1 plasmid family frequently carry multiple copies of IS-bound transposons, with the repetitive nature of the IS sequences having created a bioinformatic hurdle to reconstructing these plasmids *in silico*. The gene content of these plasmids, in particular the encoded resistance mechanisms and transposons in which they are carried have been well studied as independent elements. However there remains an incomplete understanding of the mechanisms underlying the emergence of the structurally distinct plasmid variants that are encompassed in the pSK1 family. Further, these plasmids have frequently been studied in isolation from the staphylococcal population from which they were recovered. Subsequently, the evolutionary dynamics surrounding the pSK1 plasmid family and the impact that plasmid evolution has had on the wider bacterial population is largely unknown.

4.1.1 Project Aims

The population genomics conducted in the two previous chapters of this thesis focused on the evolution of the core genome of the HA-MRSA clone ST239 in Australia and the CA-MRSA clone AK3 (ST5) in New Zealand, with the latter delving in to the structure of the autonomously mobile but chromosomally integrated *fusC*-harbouring SCC elements. Neither of these studies examined plasmids, nor their contribution to the evolution of multidrug resistant *S. aureus*. Therefore, the primary aims of this work were to explore the evolution dynamics surrounding the pSK1 plasmid family and the impact that this has had on antimicrobial resistance in ST239 MRSA. This involved using a combined genomic and phenotypic approach to: (i) reconstruct the pSK1 family plasmids circulating in this population, characterise the different structural variants and uncover the underlying mechanisms for these changes; (ii) develop a model for the evolution of this plasmid family and compare that to a model inferred from the ST239 core genome; and (iii) explore the impact that pSK1 plasmid evolution has had on phenotypic antibiotic resistance and disinfectant tolerance in the Australian ST239 MRSA population.

4.1.2 Chapter Overview

Presented in this chapter are the results of experiments conducted to address the three objectives described above, all of which have been answered in the publication presented in section 4.2. The supplementary results, methods, figures and tables from this publication are presented in section 4.2, with the supplementary dataset provided in Appendix 3. A summary of this chapter is presented in section 4.3. The implications and limitations of this work within the wider context of this thesis are discussed in Chapter 5.

4.2 Results

4.2.1 Publication: Remodelling of pSK1 family plasmids and enhanced chlorhexidine tolerance in a dominant hospital lineage of methicillin-resistant *Staphylococcus aureus*

Baines, S. L., S. O. Jensen, N. Firth, A. Gonçalves da Silva, T. Seemann, G. P. Carter, D. A. Williamson, B. P. Howden and T. P. Stinear (2019). Antimicrobial Agents and Chemotherapy. DOI: 10.1128/AAC.02356-18.

This article was selected as the Editor's Pick for Mechanisms of Resistance in this issue of Antimicrobial Agents and Chemotherapy.

Remodeling of pSK1 Family Plasmids and Enhanced Chlorhexidine Tolerance in a Dominant Hospital Lineage of Methicillin-Resistant *Staphylococcus aureus*

 Sarah L. Baines,^{a,e}  Slade O. Jensen,^b Neville Firth,^c Anders Gonçalves da Silva,^d  Torsten Seemann,^{a,d,e}
 Glen P. Carter,^{a,d}  Deborah A. Williamson,^{a,d,e}  Benjamin P. Howden,^{d,e,f}  Timothy P. Stinear^{a,e}

^aDepartment of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

^bMicrobiology and Infectious Diseases, School of Medicine, Ingham Institute for Applied Medical Research, University of Western Sydney, Sydney, New South Wales, Australia

^cSchool of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia

^dMicrobiological Diagnostic Unit Public Health Laboratory, Department of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

^eDoherty Applied Microbial Genomics, Department of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

^fInfectious Diseases Department, Austin Health, Melbourne, Victoria, Australia

ABSTRACT *Staphylococcus aureus* is a significant human pathogen whose evolution and adaptation have been shaped in part by mobile genetic elements (MGEs), facilitating the global spread of extensive antimicrobial resistance. However, our understanding of the evolutionary dynamics surrounding MGEs, in particular, how changes in the structure of multidrug resistance (MDR) plasmids may influence important staphylococcal phenotypes, is incomplete. Here, we undertook a population and functional genomics study of 212 methicillin-resistant *S. aureus* (MRSA) sequence type 239 (ST239) isolates collected over 32 years to explore the evolution of the pSK1 family of MDR plasmids, illustrating how these plasmids have coevolved with and contributed to the successful adaptation of this persistent MRSA lineage. Using complete genomes and temporal phylogenomics, we reconstructed the evolution of the pSK1 family lineage from its emergence in the late 1970s and found that multiple structural variants have arisen. Plasmid maintenance and stability were linked to IS256- and IS257-mediated chromosomal integration and disruption of the plasmid replication machinery. Overlaying genomic comparisons with phenotypic susceptibility data for gentamicin, trimethoprim, and chlorhexidine, it appeared that pSK1 has contributed to enhanced resistance in ST239 MRSA isolates through two mechanisms: (i) acquisition of plasmid-borne resistance mechanisms increasing the rates of gentamicin resistance and reduced chlorhexidine susceptibility and (ii) changes in the plasmid configuration linked with further enhancement of chlorhexidine tolerance. While the exact mechanism of enhanced tolerance remains elusive, this research has uncovered a potential evolutionary response of ST239 MRSA to biocides, one of which may contribute to the ongoing persistence and adaptation of this lineage within health care institutions.

KEYWORDS *Staphylococcus aureus*, antibiotic resistance, chlorhexidine, genomics, plasmids

Mobile genetic elements (MGEs) play a central role in microbial evolution, serving as a mechanism by which genetic material can be transferred, disseminated, and rearranged, allowing for rapid adaptation to new and changing environments. Nowhere is this more apparent than in the global dissemination of genes encoding

Citation Baines SL, Jensen SO, Firth N, Gonçalves da Silva A, Seemann T, Carter GP, Williamson DA, Howden BP, Stinear TP. 2019. Remodeling of pSK1 family plasmids and enhanced chlorhexidine tolerance in a dominant hospital lineage of methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 63:e02356-18. <https://doi.org/10.1128/AAC.02356-18>.

Copyright © 2019 Baines et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Sarah L. Baines, bainess@unimelb.edu.au, or Benjamin P. Howden, bhowden@unimelb.edu.au.

B.P.H. and T.P.S. are joint senior authors.

Received 7 November 2018

Returned for modification 9 December 2018

Accepted 8 February 2019

Accepted manuscript posted online 19 February 2019

Published 25 April 2019

mechanisms of antimicrobial resistance and virulence in populations of clinically significant bacteria (1–4). *Staphylococcus aureus* is a leading cause of bacterial infections in humans, and invasive staphylococcal disease is associated with significant morbidity and mortality (5, 6).

One of the oldest pandemic lineages of *S. aureus* is multilocus sequence type 239 (ST239), a multidrug-resistant, health care-associated (HA) methicillin-resistant *S. aureus* (MRSA) clone first identified in the late 1970s (7–9). Multiple studies have used genomics to explore the evolution of ST239 MRSA, providing insight into its global spread, extensive antimicrobial resistance repertoire, and persistence in health care environments (10–15). In Australia, ST239 has been a dominant HA-MRSA lineage for nearly 4 decades, and although its prevalence is on the decline, ST239 is still regularly identified as a cause of invasive disease in Australia, according to national surveillance reports (<http://agargroup.org.au/agar-surveys/>). We have previously described the long-term evolution of ST239 MRSA in eastern Australian hospitals to be one of convergent and adaptive evolution of two genetically distinct ST239 clades (termed the Australian and Asian-Australian clades) toward increased antimicrobial resistance at the cost of attenuated virulence (15). This initial work largely focused on changes occurring within conserved regions of the genome, with exploration of the accessory genome being limited.

Early ST239 MRSA isolates were first recognized in Australia because they displayed resistance to gentamicin (16–18), encoded by a bifunctional acetyltransferase-phosphotransferase gene [*aac(6′)-aph(2′)*] commonly carried on pSK1-like plasmids as part of the composite transposon Tn4001 (19, 20). pSK1 represents a family of intermediate-sized (20- to 40-kb), theta-replicating staphylococcal plasmids that have been recovered in Australia and the United Kingdom (19, 21–24). In addition to *aac(6′)-aph(2′)*, other antibiotic resistance elements are variably carried by these plasmids, including *blaZ* (encoding a penicillinase), carried as part of Tn552 (22, 23), and *dfrA* (encoding an insensitive dihydrofolate reductase conferring trimethoprim resistance), carried as part of Tn4003, which represents a cointegrated remnant of a pSK639-like plasmid (25, 26). Further, harbored by pSK1-like plasmids is a *qacA* gene encoding a quaternary ammonium compound (QAC) multidrug efflux pump (27), which can mediate tolerance to cationic biocides, most notably, chlorhexidine (28). An outbreak of ST239 MRSA in the United Kingdom in the early 2000s was attributed to the presence of a plasmid-borne *qacA* gene and predicted biocide tolerance (29, 30).

The divalent cationic biocide chlorhexidine digluconate (CHX) was first described in 1954 and is a fundamental component of infection control practices to prevent nosocomial infections (31, 32). It is one of the most widely used antiseptic agents because of its broad spectrum of activity against bacteria, fungi, and enveloped viruses, as well as its good safety record and general tolerability (33–35). Resistance to in-use concentrations, typically, a 0.2% to 4.0% CHX solution in water, has not been reported in *S. aureus*; however, the phenomenon of enhanced tolerance appears to be increasingly common (36, 37) and is reviewed elsewhere (38). CHX tolerance has been associated with the acquisition or mutation of biocide-active efflux pumps, with the most commonly reported being the QAC efflux systems (28, 39, 40). In *S. aureus* this is primarily *qacA* (28, 38). Most reports of staphylococcal populations evolving CHX tolerance either have demonstrated a phenotypic shift in susceptibility or have detected an increase in the prevalence of genes encoding efflux systems (38, 41). While several studies have combined phenotypic and molecular data, this work has largely been performed on clonally diverse or genomically undefined populations, and thus, the evolutionary mechanisms responsible are often unclear (36, 41).

Here we present a detailed exploration of the evolutionary dynamics surrounding the pSK1 family of plasmids, focusing on a single well-defined staphylococcal lineage, the ST239 MRSA population circulating in Australia. The primary aims of this work were to (i) describe the evolution of the pSK1 plasmid family in ST239 MRSA, with a specific focus on the emergence of plasmid structural variants, and (ii) explore the phenotypic

consequences of plasmid evolution on the ST239 population, with a particular interest in the development of enhanced CHX tolerance.

RESULTS AND DISCUSSION

An *in silico* examination of the temporally diverse, global collection of 531 ST239 isolates looking for orthologs of the pSK1 genes revealed that pSK1-like plasmids have been maintained in the ST239 population over time but are largely harbored by a single clade, the Australian clade (Fig. 1A). The earliest isolates found to harbor pSK1 were identified in 1980, consistent with when this clone first appeared in the region (16, 17, 18). This suggested the extended coevolution of the pSK1 plasmid family and the Australian ST239 clade and subsequently indicated that this population could serve as a model in which to study the evolution of multidrug resistance (MDR) staphylococcal plasmids and the impact that extended plasmid maintenance may have on antimicrobial resistance and biocide tolerance.

Prevalence of pSK1-like plasmids in ST239 MRSA. Of the 124 Australian clade isolates, 92 (74.1%) carried a pSK1-like plasmid, based on the presence of plasmid gene orthologues that were largely syntenic (Fig. S1B). A comparison of plasmid carriage with the year of isolation suggested that the prevalence of pSK1-like plasmids had increased over time (Fig. 1B). While this finding may be biased by the limited number of historical isolates available, the nearly ubiquitous presence of pSK1-like plasmids in contemporary isolates (98.4% of isolates recovered after 2005) strongly suggested an evolutionary benefit for the acquisition and maintenance of these plasmids in this population.

There was no evidence of plasmid sharing between the Australian and Asian-Australian clades (see Fig. S1 in the supplemental material). However, this finding was not surprising due to the geographic separation of the two clades; the Australian clade has been predominantly recovered in the states of New South Wales and Queensland since the early 2000s, and the newer Asian-Australian clade has almost exclusively been recovered in Victoria since its introduction (15) (Fig. 1 and S1A). Further, a large proportion of the Asian-Australian clade isolates (67/88 isolates, 76.1%) harbored a different MDR plasmid, specifically, a pTW20_1-like plasmid (29) (Fig. S2). While both pSK1 and pTW20_1 share the same mode of replication (they are theta-replicating, RepA_N plasmids [43]), it is unclear whether these plasmids are incompatible, as the nucleotide sequence homology of the replication initiation A (*repA*) genes is only 74.8%.

Structural variants of pSK1. A closer examination of the 92 isolates in which a pSK1-like plasmid was identified showed considerable variation in the presence of plasmid gene orthologs (Fig. S1B). When aligned to a phylogenetic tree of the Australian ST239 population, it became apparent that the variation observed was phylogenetically correlated and suggested the emergence or acquisition of a limited number of pSK1-like structural variants (SVs) with subsequent clonal expansion (Fig. S1). To characterize these pSK1-like SVs, representatives for the six most common orthologue patterns underwent long-read sequencing and complete genome assembly. Additionally, the short-read *de novo* assemblies of all isolates were mined for unique plasmid features to determine the prevalence of each SV in the wider ST239 population (see the Results section in the supplemental material). Using this approach, eight distinct pSK1-like plasmids were identified, having arisen largely through the actions of IS256 and/or IS257. These changes can be classified into three categories, described below.

(i) IS-mediated gain/loss of the composite transposons Tn4001 and Tn4003. The structural change consisting of the insertion sequence (IS)-mediated gain/loss of the composite transposons Tn4001 and Tn4003 was observed in three SVs, termed SV1, SV3, and SV4 (Fig. 2). The first (SV1) was found to lack the aminoglycoside resistance-conferring Tn4001 and was structurally equivalent to the previously identified pSK7 (Fig. 2A) (22, 44). With only a single copy of the IS256 target site duplications (TSD) that flank Tn4001 in pSK1 being identified, it was unclear whether SV1 had lost Tn4001 or had never acquired it. However, as an SV1-like plasmid was recovered only from historic isolates located close to the root of a time-aware phylogenetic tree for the Australian

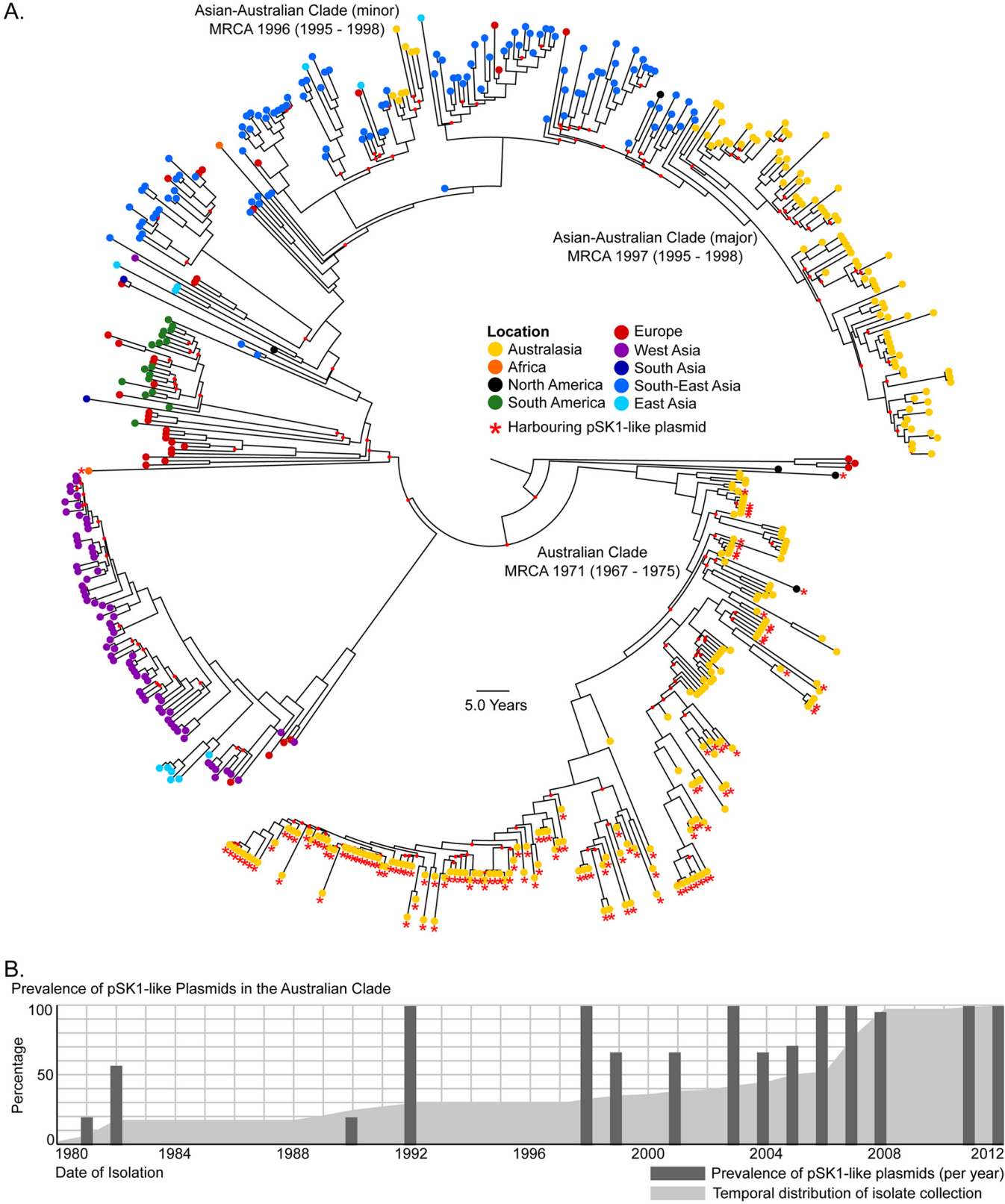
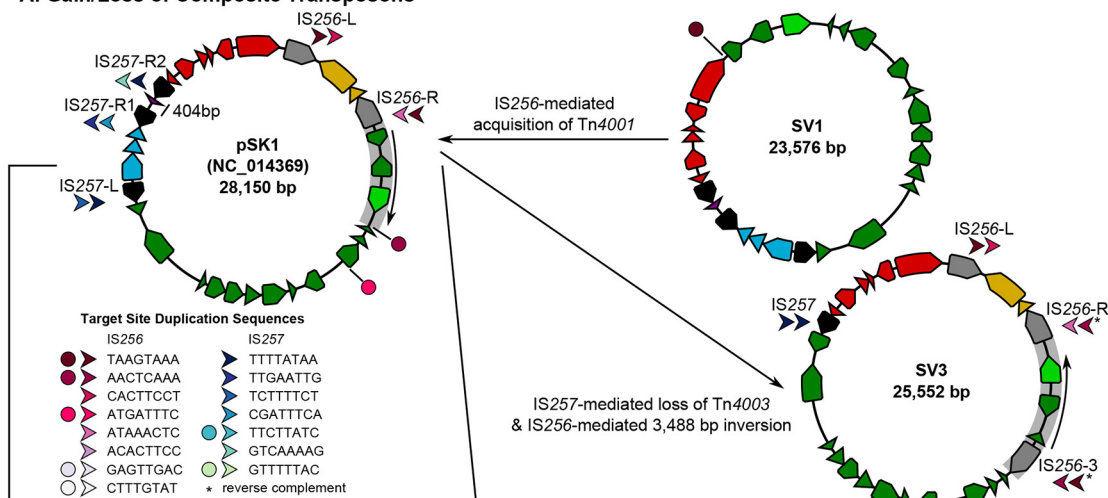


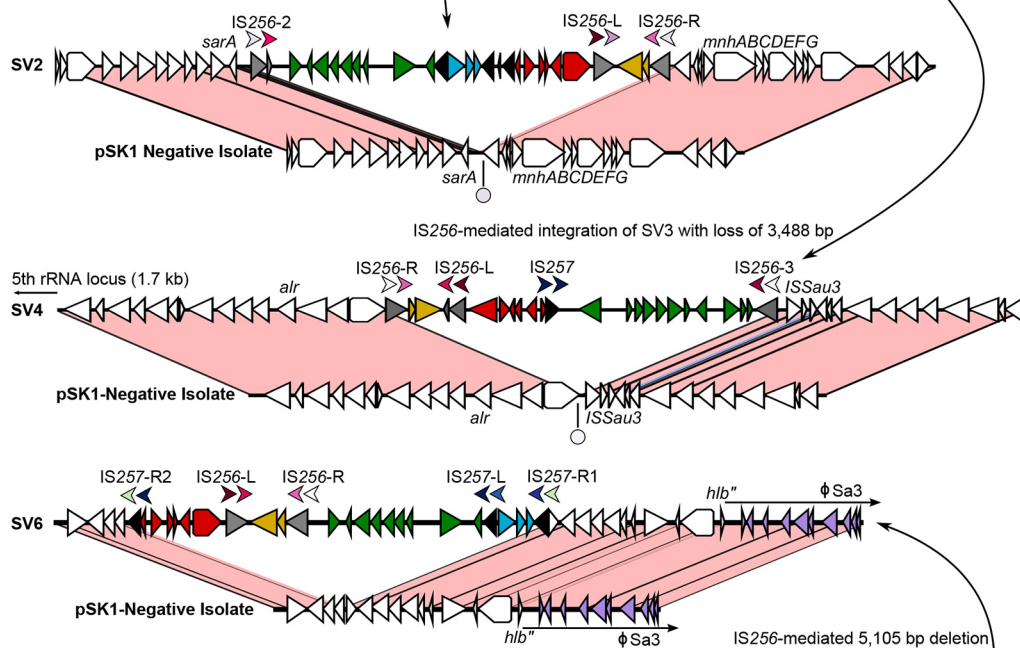
FIG 1 Global population structure of ST239 *S. aureus* isolates and prevalence of pSK1-like plasmids. (A) Illustrated is a maximum clade credibility tree inferred from the whole-genome sequence alignment of the 531 international ST239 *S. aureus* isolates. Tips are colored based on location (refer to the key). Internal nodes with <95% posterior support are indicated by a small red dot. The Australian and two Asian-Australian clades (major and minor) and estimates for the most recent common ancestor (MRCA) are indicated and are displayed as the median year (95% highest posterior density interval). The presence of a pSK1-like plasmid is indicated by a red star. (B) Temporal distribution of the Australian isolates included in the study (the cumulative total is indicated by the light gray region). The proportion of isolates harboring a pSK1-like plasmid (per year sampled) is indicated by the dark gray bars.

A. Gain/Loss of Composite Transposons



B. Chromosomal Integration & Loss of Replication Machinery

IS256-mediated integration of pSK1 with loss of 5,175 bp



C. Inversion & Disruption of Prophage Sa3

IS257-mediated integration of pSK1 and loss of 404 bp

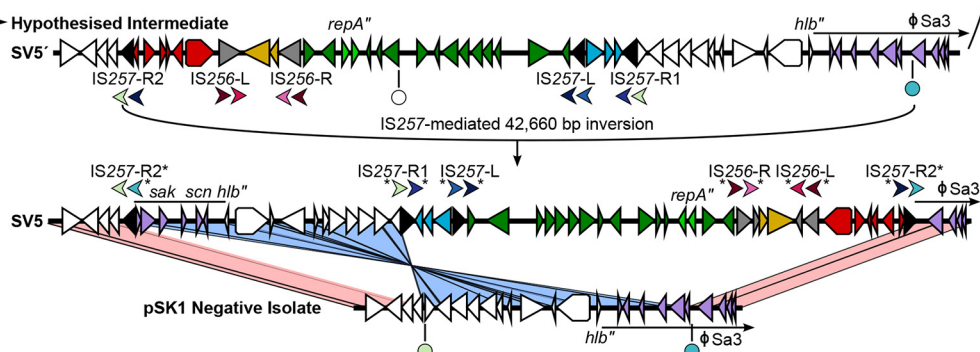


FIG 2 pSK1-like structural variants. Illustrated is a schematic of the structure and development of the pSK1-like plasmid variants. Plasmid genes are colored based on defined pSK1 regions (see Fig. S2A in the supplemental material), IS256 is gray, and IS257 is

(Continued on next page)

clade (see the Results section in the supplemental material), SV1 likely represents a progenitor of pSK1 prior to gaining Tn4001. Both SV3 and SV4 were found to lack Tn4003 (Fig. 2A and B). Instead, they contained only a single copy of IS257 with the same TSD that flanked this region in pSK1 in a configuration equivalent to that in pSK14 (22, 44). Unlike SV1, the phylogenetic location of isolates carrying SV3 and SV4 strongly supports deletion of Tn4003 (see the Results section in the supplemental material).

(ii) IS-mediated chromosomal integration and disruption of the plasmid replication machinery. Chromosomal integration was observed in five pSK1-like SVs. The first (SV2) emerged from IS256-mediated integration of pSK1 adjacent to the staphylococcal accessory regulator A (*sarA*) gene (Fig. 2B). Likewise, SV4 emerged from IS256-mediated integration of SV3 adjacent to a predicted aerobactin biosynthesis gene (Fig. 2B). In both cases, the formation of circular intermediates that excluded segments of the plasmid backbone is likely responsible for the loss of six and three syntenic genes in SV2 and SV4, respectively. The three remaining SVs had all arisen from a single IS257-mediated chromosomal integration event. Only SV5 and SV6 were represented in the complete plasmids; therefore, a hypothesized structure has been proposed for the other SV, termed SV5' (Fig. 2B and C). In this case, chromosomal integration of pSK1 had occurred 9.2 kb upstream of a disrupted β -hemolysin gene (with the disruption resulting from integration of prophage Sa3), with the formation of a shortened circular intermediate likely being responsible for the partial deletion of Tn4003 (Fig. 2C). SV5 (and likely SV5') had an additional 45-bp deletion in *repA*, specifically occurring within the Rep binding boxes, which are considered essential for a functional protein (45), and SV6 had an IS256-mediated deletion of six syntenic genes from the plasmid backbone (Fig. 2B).

(iii) IS-mediated inversion. There were two inversion events identified. The first was an IS257-mediated inversion of a 42.7-kb region encompassing SV5', giving rise to the structure termed SV5 (Fig. 2C). This inversion reversed the orientation of all SV5' genes and split ϕ Sa3 (Fig. 2C). A second smaller 3.5-kb IS256-mediated inversion event was identified in SV3, reversing the orientation of three genes, including *repA* and the plasmid partitioning (*par*) gene (Fig. 2A). This is the same region deleted in SV4. While the exact consequences of these inversions are unknown, the division and partial inversion of ϕ Sa3 should prevent future excision of the prophage.

Convergent evolution of pSK1-like plasmids. Analysis of unique plasmid features in the short-read *de novo* assemblies enabled assignment of all isolates harboring a pSK1-like plasmid to one of the eight plasmid SVs (see the Results section in the supplemental material). When overlaid onto a time-aware phylogenetic tree, these data demonstrated a clear pattern of stepwise change (Fig. 3); once a novel pSK1-like plasmid emerged, it was maintained and structurally remodeled during clonal expansion but rarely, if ever, horizontally transferred between ST239 isolates. Based on these data, we propose the following evolutionary history for the pSK1 plasmid family in the Australian ST239 population (illustrated in Fig. 3).

The first pSK1-like plasmids to have appeared in the Australian ST239 clade were SV1 and pSK1, which appeared in the late 1970s, consistent with the first identification of a gentamicin-resistant MRSA isolate in Australia (16–18). It was estimated that pSK1 emerged in 1979 (95% highest posterior density interval [HPDI], 1978 to 1980), following the acquisition of Tn4001 into the ancestral SV1 plasmid. In the following 3 decades there was the emergence of SV3 in about 1991 (HPDI, 1989 to 1994), in which Tn4003 was lost and the replication machinery was inverted. There had been at least three

FIG 2 Legend (Continued)

black. All chromosomal genes are white, and those encompassed within ϕ Sa3 are lilac. Arrows denote upstream/downstream target site duplications (TSD), and the direction denotes the IS orientation. Circles represent a single copy of a TSD not adjacent to an IS element. Arrows and circles are colored to reflect unique sequences (refer to the key), and an asterisk has been used to indicate a TSD present in the reverse complement to what was expected. In the structural comparisons, connected regions share $\geq 98\%$ nucleotide sequence identity and are colored pink or blue to indicate the matching or reverse orientation, respectively. (A) Emergence of pSK1-like variants through IS-mediated loss/gain of the composite transposons; (B) chromosomal integration with loss/disruption of the plasmid replication machinery; (C) inversion with fragmentation of ϕ Sa3.

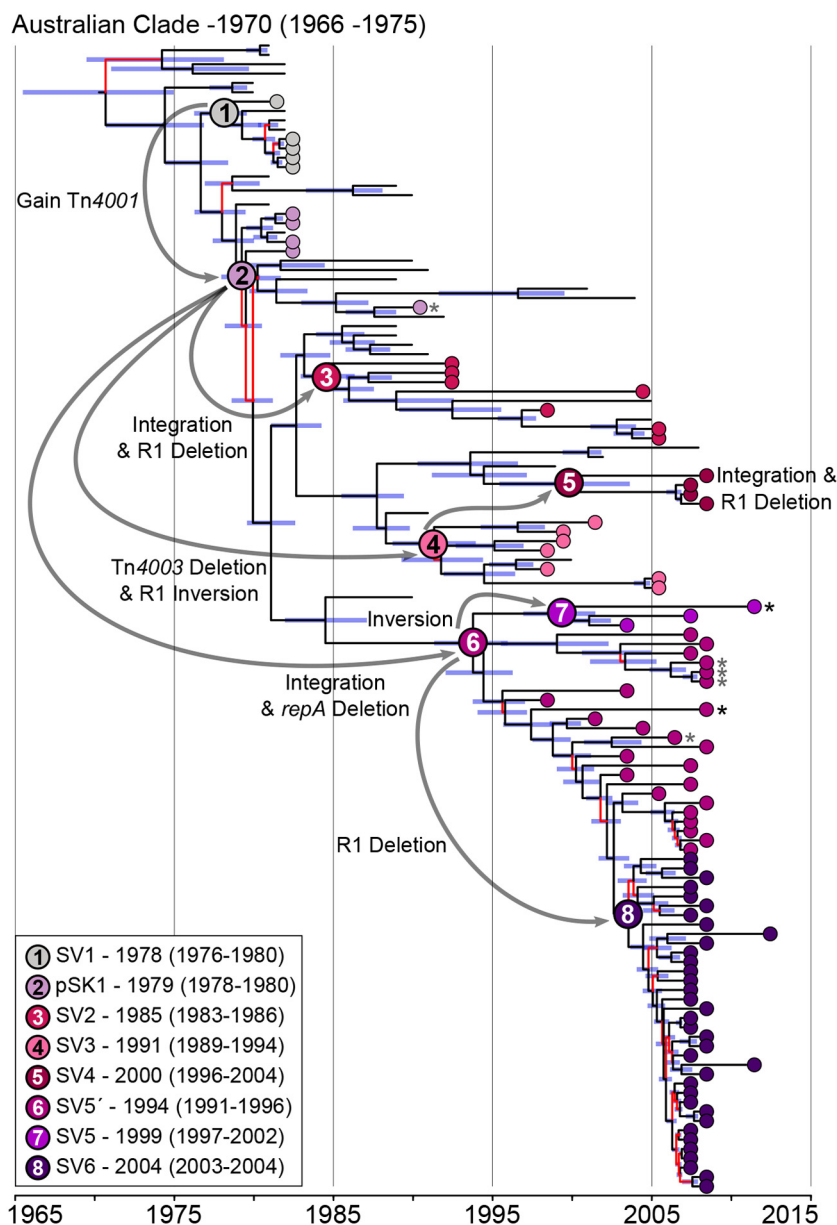


FIG 3 Evolutionary history of pSK1-like plasmids in the Australian ST239 clade. Illustrated is a maximum clade credibility tree inferred from the whole-genome sequence alignment of the Australian clade ($n = 124$). Branches with $<95\%$ posterior support are colored red, and the 95% highest posterior density interval for node heights is represented by the blue bars. Isolates identified as harboring a pSK1-like plasmid are indicated by a circle located at the branch tip and colored based on the SV identified (refer to the key). The most recent common ancestor for each SV is indicated by the larger node circle, with black numbers indicating an extrachromosomal SV and white numbers indicating a chromosomally integrated SV. Temporal estimates for these nodes have been provided and are displayed as SV-median year (95% highest posterior density interval). Gray arrows illustrate the likely order and type of structural change that has occurred during the evolution of the pSK1-like plasmid population. Isolates which demonstrated a plasmid gene orthologue pattern suggestive of further structural changes are indicated by an asterisk (black for deletion of the plasmid replication machinery and gray for the loss of Tn4003).

chromosomal integration events, resulting in the emergence of SV2 in about 1985 (HPDI, 1983 to 1986), SV5' in about 1994 (HPDI, 1991 to 1996), and SV4 in about 2000 (HPDI, 1996 to 2004). These events likely involved pSK1 as the ancestral plasmid or SV3 as the immediate ancestor of SV4. Inversion of SV5' in 1999 (HPDI, 1997 to 2002) led to the emergence of SV5.

In addition to this stepwise change, there is also evidence for convergent evolution in the plasmid population. The three chromosomal integration events all occurred

independently. Examination of all 92 isolates harboring a pSK1-like plasmid identified a further two independent deletion events (Fig. 3). There is a clear role for chromosomal integration in improving the maintenance of plasmid genes in a population, and the subsequent deletion or disruption of the plasmid replication machinery is needed for stability of the integrant as it removes interference with chromosomal replication (46). In the Australian clade, all but two isolates that descended from an ancestral genome with an integrated pSK1-like structure have maintained it (Fig. 3). These data strongly suggest that the emergence of these phylogenetically distinct but structurally similar SVs is the result of a significant but unknown evolutionary pressure acting on the plasmid population. It is plausible that the introduction and/or increased use of antimicrobial agents and disinfectants, specifically, those for which mechanisms of resistance or tolerance are carried by these plasmids, could be contributing to the evolutionary pressure promoting plasmid maintenance and driving chromosomal integration.

Impact of plasmid evolution on antimicrobial resistance and disinfectant tolerance. To explore the impact of plasmid evolution on antimicrobial resistance and disinfectant tolerance, 211 isolates underwent testing for susceptibility to gentamicin, trimethoprim, and chlorhexidine. Isolates representing both the Australian and Asia-Australian clades were tested to enable comparisons between the two clades and discern which trends may be associated with the acquisition and/or evolution of the pSK1-like plasmids.

(i) Gentamicin. Phenotypic resistance to gentamicin was detected in 185 (87.7%) isolates. Three acquired genes encoding aminoglycoside-modifying enzymes (AMEs) were identified (47) (Table 1): (i) the bifunctional *aac(6')-aph(2'')*, found in Tn4001 and other chromosomal and phage locations (48–50); (ii) an adenyltransferase gene [*aadD*, also known as ANT(4')-Ia] (51, 52); and (iii) a phosphotransferase gene [*aph(3')-IIIa*] (53), which has previously been found to be colocated with *aac(6')-aph(2'')* in an integrated φ SP β -like phage in ST239 isolates closely related to those in the Asian-Australian clade (15, 29). The bifunctional AME was common among both the Australian and the Asian-Australian clades (79.7% and 93.2%, respectively). Conversely, the distribution of the monofunctional AMEs was more distinct (Table 1). Phenotypically, the Asian-Australian clade demonstrated a significantly higher average MIC of gentamicin than the Australian clade ($P < 0.0001$; Table S1). However, as most isolates harbored multiple AMEs, phenotypic resistance could not be attributed to a single gene. With that being said, the distribution of these genes suggested that *aph(3')-IIIa* is likely responsible for the high-level resistance observed in the Asian-Australian clade, with *aac(6')-aph(2'')* and *aadD* contributing only low-level resistance in the Australian clade (Table S1).

To further investigate this phenotype, linear models were developed to look for temporal trends in the data. These models identified a significant increase in the median gentamicin MIC for the ST239 population over time ($P < 0.001$). However, this trend was observed only in the Australian clade when the two clades were modeled separately (Fig. S3). This finding is similar to what has previously been observed in the ST239 population with reduced glycopeptide and daptomycin susceptibility, hypothesized to be the result of two evolutionary phenomena: (i) the introduction of the more resistant Asian-Australian clade into the region with successful local expansion and (ii) adaptive evolution within the Australian clade, collectively shifting the phenotype of the population over time (15). These same phenomena are likely contributing here. The Asian-Australian clade had already developed high-level gentamicin resistance, prior to arriving in Australia, through the acquisition of an φ SP β -like phage carrying both *aac(6')-aph(2'')* and *aph(3')-IIIa*. Concurrently, the Australian clade had undergone adaptive evolution, with the acquisition of *aac(6')-aph(2'')* and/or *aadD* contributing to a significant increase in the MIC (Table S1) and with the uptake of a pSK1-like plasmid serving as one mechanism by which *aac(6')-aph(2'')* could have been acquired. Within the pSK1-like plasmid-harboring population, there were no discernible temporal trends

TABLE 1 Population distributions of antimicrobial and biocide resistance genes and phenotypic resistance profiles^a

Isolate	Median (range)		MIC (mg/liter)	Median (range)		Acquired resistance gene(s) ^c (no. of isolates)			
	Gentamicin ^a	Trimethoprim ^a		Chlorhexidine ^b	MBC (mg/liter) ^b	AME genes	dfr genes	qac genes	
All ST239 isolates (n = 211)	32 (0.023–256)	>32		3 (1–6)	6 (2–16)	<i>aac(6′)-aph2″</i> (180), <i>aadD</i> (72), <i>aph(3′)-III</i> (95)	<i>dfrA</i> (76), <i>dfrG</i> (88)	<i>qacA</i> (156), <i>qacC</i> (2)	
Asian-Australian clade (n = 88)	256 (0.38–256)			3 (1.5–4)	6 (3–12)	<i>aac(6′)-aph2″</i> (82), <i>aph(3′)-III</i> (87)	<i>dfrG</i> (88)	<i>qacA</i> (67)	
Australian clade (n = 123)	16 (0.023–256)			4 (1–6)	8 (2–16)	<i>aac(6′)-aph2″</i> (98), <i>aadD</i> (72), <i>aph3′-III</i> (8)	<i>dfrA</i> (76)	<i>qacA</i> (89), <i>qacC</i> (2)	
pSK1 plasmid (n = 91)	24 (0.032–256)			4 (1–6)	8 (2–16)	<i>aac(6′)-aph2″</i> (85), <i>aadD</i> (67), <i>aph3′-III</i> (5)	<i>dfrA</i> (76)	<i>qacA</i> (89)	

^aPhenotypic susceptibility to gentamicin and trimethoprim was determined by Etest.
^bPhenotypic susceptibility to chlorhexidine was determined by broth microdilution.
^cIdentification of resistance genes was determined by local alignment; a minimum alignment of 70% of the gene length and >95% nucleotide sequence identity were required to call a match.
^dAbbreviations: AME, aminoglycoside-modifying enzyme; MBC, minimum bactericidal concentration.

TABLE 2 Investigation of chlorhexidine tolerance in Australian ST239 MRSA^c

Isolate population	Group 1		Group 2		P values for MICs, MBCs in group 1 vs group 2
	Characteristic (no. of isolates)	Median MIC, MBC (mg/liter)	Characteristic (no. of isolates)	Median MIC, MBC (mg/liter)	
All ST239 (211)	<i>qacA</i> negative (55)	0.90, 2.30	<i>qacA</i> positive (156)	3.70, 8.00	<0.0001, <0.0001
	Australian clade (123)	3.40, 7.30	Asian-Australian clade (88)	2.30, 6.70	<0.0001, NS
	pSK1-like plasmid (91)	4.20, 8.60	pTW20_1-like plasmid (67)	3.00, 7.20	<0.0001, 0.0021
Asian-Australian clade (88)	<i>qacA</i> negative (21)	1.90, 5.20	<i>qacA</i> positive (67)	3.00, 7.20	<0.0001, <0.0001
Australian clade (123)	<i>qacA</i> negative (34)	1.80, 4.60	<i>qacA</i> positive (89)	4.30, 8.60	<0.0001, <0.0001
	Extrachromosomal SVs (17)	3.20, 7.80	Integrated SVs (74)	4.50, 8.74	0.0086, NS
	Integrated SVs + <i>repA</i> deletion (25) ^a	3.70, 8.50	Integrated SV + MG deletion (49) ^b	4.70, 8.60	NS, NS

^aThe population includes SV5' and SV5.^bThe population includes SV2, SV4, and SV6.^cAbbreviations: MBC, minimum bactericidal concentration; MG, multigene; SV, structural variant; NS, not significant.

(Fig. S3), which suggested that the emergence of the pSK1-like SVs has had no additional impact on the gentamicin MIC.

(ii) Trimethoprim. Phenotypic resistance to trimethoprim was detected in all 211 ST239 isolates, for which the MIC was >32 mg/liter. Genotypically, 75/123 isolates (61.0%) of the Australian clade and all 88 isolates of the Asian-Australian clade were found to harbor a mutated or acquired *dfp* gene (Table 1). A likely mechanism of resistance could not be identified in the remaining 48 isolates. The lack of phenotypic variation detected suggested that both the acquisition and changes in the configuration of the pSK1-like plasmids, including the deletion or alteration of Tn4003, had not impacted trimethoprim resistance due to an unidentified genotypic redundancy.

(iii) Chlorhexidine. Phenotypic tolerance to CHX, defined as an MIC of >2 mg/liter, was detected in 150 (71.1%) isolates (38). A total of 156 (73.9%) isolates were found to harbor *qacA*, and 2 were found to harbor *qacC* (Table 1). QacC is not active against CHX (38). The presence of *qacA* was significantly associated with an elevated MIC and minimum bactericidal concentration (MBC) of CHX ($P < 0.0001$) and was invariably associated with carriage of a pSK1-like plasmid (Australian clade) or a pTW20_1-like plasmid (Asian-Australian clade) (Table 2). Although *qacA* was equally prevalent in both the Australian and Asian-Australian clades (72.4% and 76.1%, respectively), the former population demonstrated a significantly higher average MIC ($P < 0.0001$, Table 2). This phenotypic disparity remained when isolates that did not carry either plasmid were excluded from the comparison. Further, no mutations in *qacAR* or variations in gene copy number that could explain this phenotypic variation were detected (see the Results section in the supplemental material), suggesting that genotypic differences occurring outside of *qacAR* were responsible.

As with gentamicin, linear models were developed to investigate temporal trends in CHX susceptibility (Fig. S3). These models identified a significant increase in the median CHX MIC and MBC of the ST239 population over time ($P < 0.001$). Again, this trend was observed in the Australian clade but not the Asian-Australian clade when modeled separately (Fig. S3). Further, this trend remained when only the pSK1-like plasmid-harboring population was modeled. This suggested that, in contrast to the evolutionary phenomena facilitating the development of reduced antibiotic susceptibility, adaptive evolution in the Australian clade appeared to be the primary contributor to enhanced CHX tolerance in the ST239 population. Comparison of the average MIC and MBC between the pSK1-like plasmids suggested that the more recently emerged SVs (SV2 to SV6) may be associated with enhanced tolerance (Fig. 4; Table S2). When isolates were grouped based on their carriage of structurally similar SVs, chromosomal integration was significantly associated with an increased CHX MIC ($P = 0.0086$; Fig. 4; Table 2). Furthermore, a multigene deletion in the plasmid backbone appeared to be associated with a further increase in the MIC, although this was statistically nonsignificant (Fig. 4; Table 2). It is unclear why these structural changes were associated only with a shift in the MIC, with all groups having demonstrated highly consistent average MBC values

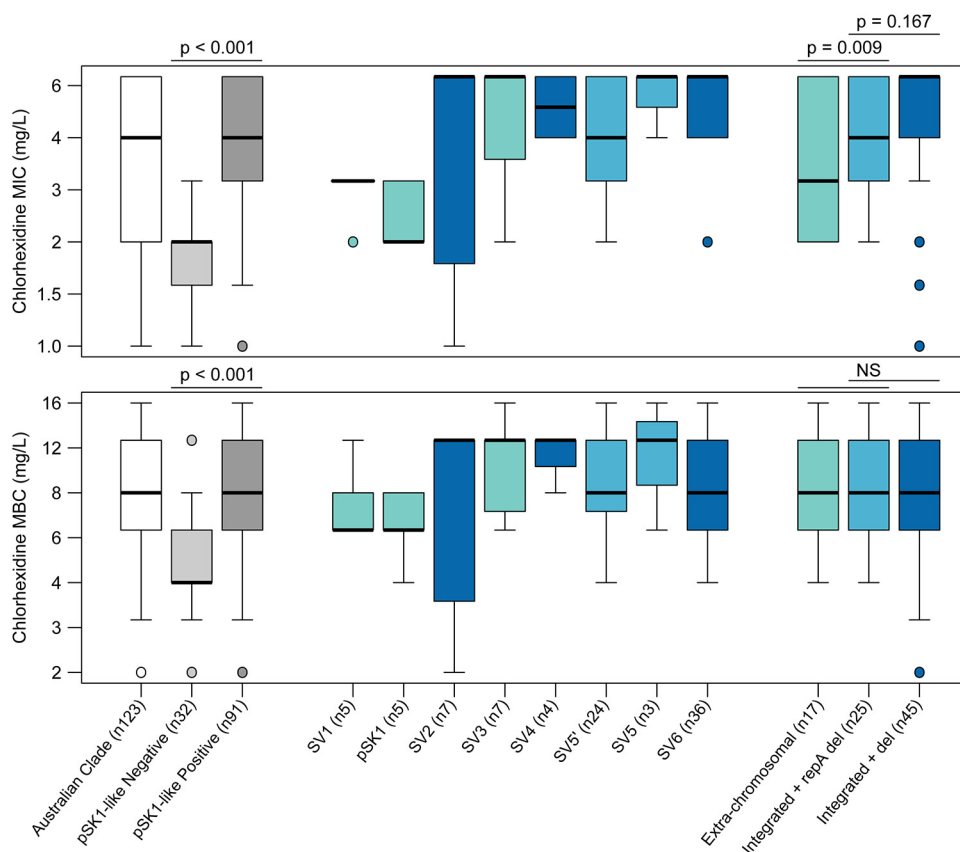


FIG 4 Phenotypic variation in chlorhexidine tolerance. Graphs illustrate the distribution of chlorhexidine MIC (top) and MBC (bottom) values in the Australian clade. Box plot features represent the population median (central black line), upper and lower quartiles (box), and range (bars), excluding outliers (circles). Box plots representing SVs are colored to reflect a plasmid structural feature: an extrachromosomal plasmid (teal) and a chromosomally integrated plasmid with either an internal *repA* deletion (del) (light blue) or a multigene deletion (dark blue).

(Fig. 4; Table S2), but this finding may reflect the specific mechanisms mediating enhanced CHX tolerance or the method utilized to measure susceptibility.

Independent association of CHX tolerance with pSK1-like plasmid evolution. To examine the strength of the association between plasmid evolution and the development of enhanced CHX tolerance, in addition to exploring other possible mechanisms that may be contributing to this phenotype, three statistical genomic techniques were utilized. A series of genome-wide association studies (GWAS) and discriminant analysis of principal components (DAPC) studies was performed, but none identified any significantly associated genotypes, outside of the genotype resulting from the presence of *qacA*-harboring plasmids, that could be responsible for enhanced CHX tolerance (see the Results section in the supplemental material). We therefore performed a modified Bayesian phylogeographic analysis, modeling the CHX phenotype rather than the geographic location of the ancestral genomes. The hypothesis was that if the development of CHX tolerance was a consequence of plasmid evolution, then the ancestral nodes (ANs) from which the SVs were estimated to have emerged should correlate with those in which a shift in MIC is predicted (Fig. 5).

In this model, a shift in MIC was predicted at the expected ANs or at the directly preceding node for the emergence of the newer SVs (SV3, SV4, SV5', SV5, and SV6; Fig. 5). In the older SVs, a shift in MIC was modeled to have occurred a few nodes following the emergence of the SV (SV1, pSK1, and SV2; Fig. 5). This discrepancy is likely due to the higher proportion of isolates not harboring a pSK1-like plasmid in close proximity to these ANs, with a shift in MIC being predicted at the ANs for subclades of isolates harboring these older SVs (Fig. 5). Overall, this model did suggest that plasmid

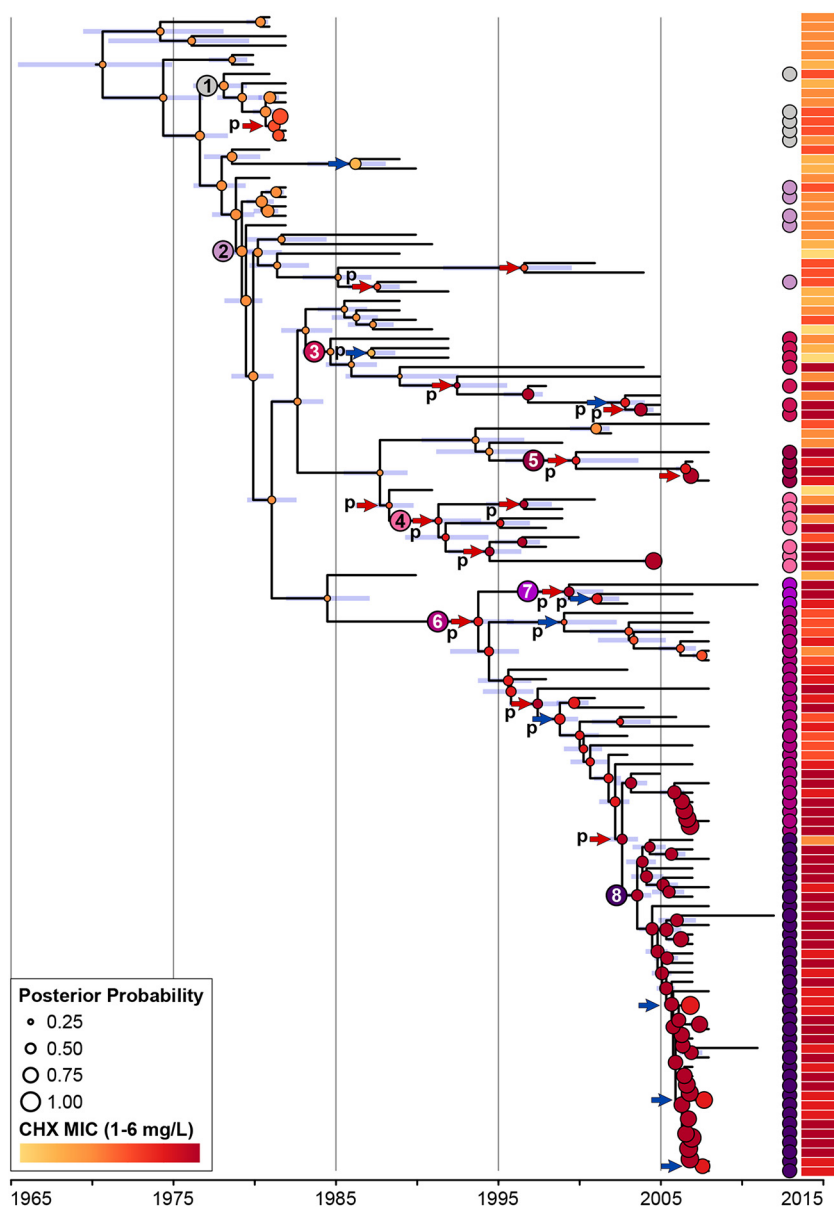


FIG 5 Bayesian phylogenetic model associating chlorhexidine tolerance with pSK1-like plasmid evolution. Illustrated is a maximum clade credibility tree inferred from the whole-genome sequence alignment of the Australian clade ($n = 124$). Isolates identified as harboring a pSK1-like plasmid are indicated by a circle located adjacent to the tree and colored based on the SV identified. Blue bars represent the 95% highest posterior density interval for the node heights. The ancestral nodes in which each SV is estimated to have emerged are indicated by a number (the numbers are the same as the numbers in Fig. 3). The estimated CHX MIC for all ancestral nodes is indicated by a circle, colored based on the MIC value and sized according to the posterior probability for the estimate (refer to the key). The 24 nodes with a predicted shift in the MIC are indicated by an arrow, colored blue or red for a decrease or increase in the MIC compared to that in the preceding node, respectively. Those that have been tentatively associated with the changes in plasmid configuration are labeled with "p." The aligned heat map illustrates the phenotypic MIC values attained for each isolate.

evolution was associated with shifts in CHX tolerance, but it is unlikely the sole contributing mechanism. Of the 24 nodes (19.5% of all nodes) in which a shift in MIC was detected, only 18 appeared to be associated with a change in plasmid configuration. Further lab-based experimentation is required to confirm these findings.

Chlorhexidine is one of the most commonly used biocides in the health care environment and a fundamental component of infection control initiatives shown to be effective in reducing the rates of invasive staphylococcal disease (54, 55). Therefore, it

is plausible that its widespread use may be contributing to the evolutionary pressures promoting maintenance of the pSK1-like plasmid family in ST239 MRSA. However, the *qacA*-encoded efflux pump is active against a range of organic cations (28, 38), and this detected increase in tolerance may extend to other biocides, making them potential drivers of plasmid evolution. In a separate finding, all pSK1-like plasmids maintained an ~9.3-kb segment of the plasmid backbone, demonstrating >95% nucleotide sequence identity to *S. warneri* plasmid pPI-1 (GenBank accession number [AB125341.3](#)) (44, 56). The genes encompassed within this region encoded protein products predicted to be cell envelope associated and involved in membrane transport and, potentially, iron acquisition, with one gene encoding a putative Fst-like toxin as part of a type I toxin-antitoxin system (44). The ubiquitous conservation of this region in the pSK1-like plasmid population suggested a potential benefit of these encoded products to the ST239 population and the possibility that they also contribute to the evolution and maintenance of this plasmid population (44).

Conclusions. Plasmids and other MGEs play a central role in the successful evolution and adaptation of bacterial populations but are often overlooked because of the challenges with examining some plasmid DNA sequences using short-read data. Here, we have provided a comprehensive analysis of the evolution of the pSK1-like plasmid population that has coevolved with the Australian ST239 MRSA lineage over multiple decades. Within this ST239 population, there are at least eight structurally distinct pSK1-like plasmids, having arisen largely through the actions of IS256 and IS257. Further, this plasmid population has undergone convergent evolution, with the repeated emergence of chromosomally integrated SVs. In investigating how plasmid evolution may alter phenotypic susceptibility to antibiotics and biocides, it was identified that plasmid acquisition is associated with a significant increase in gentamicin resistance and CHX tolerance. Further, carriage of a chromosomally integrated plasmid variant was associated with a further increase in CHX tolerance. While the mechanism mediating tolerance remains unclear, we speculate that it is linked to altered regulation of the *qacAR* efflux system due to movement of the IS elements. Collectively, these findings support the idea that the widespread and increasing use of cationic biocides may potentially be contributing to the ongoing evolution of the pSK1-like plasmid population. Although the levels of CHX at which isolates show reduced susceptibility observed in this study remain well below the in-use concentrations of CHX, they do illustrate an evolutionary response in ST239 MRSA, one that may provide an adaptive advantage for this lineage in health care institutions.

MATERIALS AND METHODS

Bacterial isolates. This study utilized a temporal (1980 and 2012) and geographically diverse collection of 212 Australian ST239 *S. aureus* isolates. Two isolates (including reference strain JKD6008) were recovered in New Zealand (57). All isolates represented cases of clinical infection. To establish the global phylogenetic context for the collection, we supplemented these data with the WGS data for a further 319 international ST239 *S. aureus* isolates. Further information about the selection of isolates and relevant demographic information can be found in the Materials section in the supplemental material.

WGS and sequence data. The whole-genome sequencing (WGS) data for 368/531 isolates (73/212 Australian isolates) have been previously published. Seven isolates were subjected to long-read sequencing to enable complete genome assembly. Information about the generation of novel sequence data and relevant WGS information can be found in the Materials section in the supplemental material.

Bioinformatic analysis. The bioinformatic analyses performed for this study have been explained in detail in the Materials section in the supplemental material and are briefly outlined here. Short-read sequence data were mapped to reference strain *S. aureus* JKD6008 (GenBank accession number [CP002120](#) [58]) using Snippy (v3.2) (<https://github.com/tseemann/snippy>). Maximum likelihood phylogenetic trees were generated with IQ-TREE (v1.6.1) (59), and maximum clade credibility trees were generated with BEAST (v2.4.7) (60). Trees were visualized in FigTree (v1.4.3) (<http://tree.bio.ed.ac.uk/software/figtree/>), and figures were assembled in Inkscape (v0.91) (<https://inkscape.org/>). Short-read sequence data were *de novo* assembled using SPAdes (v3.11.0) (61) and annotated with Prokka (v1.12) (62). Long-read sequence data were *de novo* assembled using the SMRT analysis system (v2.3.0.140936) (Pacific Biosciences), circularized and reoriented in Geneious (v8.1.5) (Biomatters), and polished with Snippy (v3.2). Plasmid structural comparisons were conducted using the Artemis Comparison Tool (63). Ortholog clustering was performed using Roary (64).

Antimicrobial susceptibility testing. Testing for phenotypic susceptibility to trimethoprim and gentamicin was performed using Etests (bioMérieux), and the results were interpreted using CLSI

guidelines (65). Susceptibility to CHX was performed using a modified broth microdilution method (see the Materials section in the supplemental material). All isolates were tested in biological triplicate, and the median values were used for statistical analysis, performed in R (v3.4.2) software (<http://www.R-project.org/>).

Accession number(s). All sequence data and complete genomes novel to this study have been submitted to the European Nucleotide Archive under project accession number [PRJEB29322](https://www.ebi.ac.uk/ena/record/PRJEB29322).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AAC.02356-18>.

SUPPLEMENTAL FILE 1, PDF file, 1.3 MB.

ACKNOWLEDGMENTS

This work was funded by the National Health and Medical Research Council, Australia (fellowship GNT1105905, Centres of Research Excellence grant GNT1102962 to B.P.H., and project GNT457454 to N.F. and S.O.J.). S.L.B. was supported by an Australian government research training program scholarship and a Victorian fellowship provided by the Victoria State government.

We thank Eloise Alison for her assistance in generating the phenotypic susceptibility data.

REFERENCES

- Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304. <https://doi.org/10.1038/35012500>.
- Frost LS, Leplae R, Summers AO, Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 3:722–732. <https://doi.org/10.1038/nrmicro1235>.
- Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. *Nat Rev Genet* 16:472–482. <https://doi.org/10.1038/nrg3962>.
- Partridge SR, Kwong SM, Firth N, Jensen SO. 2018. Mobile genetic elements associated with antimicrobial resistance. *Clin Microbiol Rev* 31:e00088-17. <https://doi.org/10.1128/CMR.00088-17>.
- van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB. 2012. Predictors of mortality in *Staphylococcus aureus* bacteremia. *Clin Microbiol Rev* 25:362–386. <https://doi.org/10.1128/CMR.05022-11>.
- Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. 2015. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* 28:603–661. <https://doi.org/10.1128/CMR.00134-14>.
- Pavillard R, Harvey K, Douglas D, Hewstone A, Andrew J, Collopy B, Asche V, Carson P, Davidson A, Gilbert G, Spicer J, Tosoloni F. 1982. Epidemic of hospital-acquired infection due to methicillin-resistant *Staphylococcus aureus* in major Victorian hospitals. *Med J Aust* 29:451–454.
- Cookson BD, Phillips I. 1988. Epidemic methicillin-resistant *Staphylococcus aureus*. *J Antimicrob Chemother* 21:57–65. https://doi.org/10.1093/jac/21.suppl_C.57.
- Dubin DT, Chikramane SG, Inglis B, Matthews PR, Stewart PR. 1992. Physical mapping of the *mec* region of an Australian methicillin-resistant *Staphylococcus aureus* lineage and a closely related American strain. *J Gen Microbiol* 138:169–180. <https://doi.org/10.1099/00221287-138-1-169>.
- Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327:469–474. <https://doi.org/10.1126/science.1182395>.
- Castillo-Ramírez S, Corander J, Marttinen P, Aldehawi M, Hanage WP, Westh H, Boye K, Gulay Z, Bentley SD, Parkhill J, Holden MT, Feil EJ. 2012. Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol* 13:R126. <https://doi.org/10.1186/gb-2012-13-12-r126>.
- Gray RR, Tatem AJ, Johnson JA, Alekseyenko AV, Pybus OG, Suchard MA, Salemi M. 2011. Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a Bayesian framework. *Mol Biol Evol* 28:1593–1603. <https://doi.org/10.1093/molbev/msq319>.
- Hsu LY, Harris SR, Chlebowicz MA, Lindsay JA, Koh TH, Krishnan P, Tan TY, Hon PY, Grubb WB, Bentley SD, Parkhill J, Peacock SJ, Holden MT. 2015. Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol* 16:81. <https://doi.org/10.1186/s13059-015-0643-z>.
- Tong SY, Holden MT, Nickerson EK, Cooper BS, Koser CU, Cori A, Jombart T, Cauchemez S, Fraser C, Wuthiekanun V, Thaipadungpanit J, Hongsuwan M, Day NP, Limmathurotsakul D, Parkhill J, Peacock SJ. 2015. Genome sequencing defines phylogeny and spread of methicillin-resistant *Staphylococcus aureus* in a high transmission setting. *Genome Res* 25:111–118. <https://doi.org/10.1101/gr.174730.114>.
- Baines SL, Holt KE, Schultz MB, Seemann T, Howden BO, Jensen SO, van Hal SJ, Coombs GW, Firth N, Powell DR, Stinear TP, Howden BP. 2015. Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*. *mBio* 6:e00080-15. <https://doi.org/10.1128/mBio.00080-15>.
- Graham DR, King K, Brady LM, Harkness JL. 1981. Gentamicin-resistant staphylococci. *Lancet* 318:698–699. [https://doi.org/10.1016/S0140-6736\(81\)91035-7](https://doi.org/10.1016/S0140-6736(81)91035-7).
- Gedney J, Lacey RW. 1982. Properties of methicillin-resistant staphylococci now endemic in Australia. *Med J Aust* 1:448–450.
- Gilbert GL, Asche V, Hewstone AS, Mathiesen JL. 1982. Methicillin-resistant *Staphylococcus aureus* in neonatal nurseries. Two years' experience in special-care nurseries in Melbourne. *Med J Aust* 1:455–459.
- Gillespie MT, May JW, Skurray RA. 1984. Antibiotic susceptibilities and plasmid profiles of nosocomial methicillin-resistant *Staphylococcus aureus*: a retrospective study. *J Med Microbiol* 17:295–310. <https://doi.org/10.1099/00222615-17-3-295>.
- Lyon BR, May JW, Skurray RA. 1984. Tn4001: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Mol Gen Genet* 193:554–556. <https://doi.org/10.1007/BF00382099>.
- Lyon BR, Iuorio JL, May JW, Skurray RA. 1984. Molecular epidemiology of multiresistant *Staphylococcus aureus* in Australian hospitals. *J Med Microbiol* 17:79–89. <https://doi.org/10.1099/00222615-17-1-79>.
- Skurray RA, Rouch DA, Lyon BR, Gillespie MT, Tennent JM, Byrne ME, Messerotti LJ, May JW. 1988. Multiresistant *Staphylococcus aureus*: genetics and evolution of epidemic Australian strains. *J Antimicrob Chemother* 21:19–39. https://doi.org/10.1093/jac/21.suppl_C.19.
- Wright C, Byrne M, Firth N, Skurray R. 1998. A retrospective molecular analysis of gentamicin resistance in *Staphylococcus aureus* strains from UK hospitals. *J Med Microbiol* 47:173–178. <https://doi.org/10.1099/00222615-47-2-173>.
- Townsend DE, Ashdown N, Bolton S, Bradley J, Duckworth G, Moorhouse EC, Grubb WB. 1987. The international spread of methicillin-resistant

- Staphylococcus aureus*. J Hosp Infect 9:60–71. [https://doi.org/10.1016/0195-6701\(87\)90097-1](https://doi.org/10.1016/0195-6701(87)90097-1).
25. Rouch DA, Messerotti LJ, Loo LS, Jackson CA, Skurray RA. 1989. Trimethoprim resistance transposon Tn4003 from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS257. Mol Microbiol 3:161–175. <https://doi.org/10.1111/j.1365-2958.1989.tb01805.x>.
 26. Firth N, Skurray RA. 1998. Mobile elements in the evolution and spread of multiple-drug resistance in staphylococci. Drug Resist Updat 1:49–58. [https://doi.org/10.1016/S1368-7646\(98\)80214-8](https://doi.org/10.1016/S1368-7646(98)80214-8).
 27. Tennent JM, Lyon BR, Midgley M, Jones IG, Purewal AS, Skurray RA. 1989. Physical and biochemical characterization of the *qacA* gene encoding antiseptic and disinfectant resistance in *Staphylococcus aureus*. J Gen Microbiol 135:1–10. <https://doi.org/10.1099/00221287-135-1-1>.
 28. Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA. 1996. Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity. Proc Natl Acad Sci U S A 93:3630–3635. <https://doi.org/10.1073/pnas.93.8.3630>.
 29. Holden MTG, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, Batra R, Parkhill J, Bentley SD, Edgeworth JD. 2010. Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). J Bacteriol 192:888–892. <https://doi.org/10.1128/JB.01255-09>.
 30. Batra R, Cooper BS, Whiteley C, Patel AK, Wyncoll D, Edgeworth JD. 2010. Efficacy and limitation of a chlorhexidine-based decolonization strategy in preventing transmission of methicillin-resistant *Staphylococcus aureus* in an intensive care unit. Clin Infect Dis 50:210–217. <https://doi.org/10.1086/648717>.
 31. Septimus EJ, Schweizer ML. 2016. Decolonization in prevention of health care-associated infections. Clin Microbiol Rev 29:201–222. <https://doi.org/10.1128/CMR.00049-15>.
 32. Davies GE, Francis J, Martin AR, Rose FL, Swain G. 1954. 1:6-Di-4'-chlorophenylidiguandohexane (Hibitane). Laboratory investigation of a new antibacterial agent of high potency. Br J Pharmacol Chemother 9:192–196. <https://doi.org/10.1111/j.1476-5381.1954.tb00840.x>.
 33. Hugo WB, Longworth AR. 1964. Some aspects of the mode of action of chlorhexidine. J Pharm Pharmacol 16:655–662. <https://doi.org/10.1111/j.2042-7158.1964.tb07384.x>.
 34. Russell AD. 1986. Chlorhexidine: antibacterial action and bacterial resistance. Infection 14:212–215. <https://doi.org/10.1007/BF01644264>.
 35. Kampf G, Kramer A. 2004. Epidemiologic background of hand hygiene and evaluation of the most important agents for scrubs and rubs. Clin Microbiol Rev 17:863–893. <https://doi.org/10.1128/CMR.17.4.863-893.2004>.
 36. Hardy K, Sunnucks K, Gil H, Shabir S, Trampari E, Hawkey P, Webber M. 2018. Increased usage of antiseptics is associated with reduced susceptibility in clinical isolates of *Staphylococcus aureus*. mBio 9:e00894-18. <https://doi.org/10.1128/mBio.00894-18>.
 37. Hayden MK, Lolans K, Haffenreffer K, Avery TR, Kleinman K, Li H, Kaganov RE, Lankiewicz J, Moody J, Septimus E, Weinstein RA, Hickok J, Jernigan J, Perlin JB, Platt R, Huang SS. 2016. Chlorhexidine and mupirocin susceptibility of methicillin-resistant *Staphylococcus aureus* isolates in the REDUCE-MRSA trial. J Clin Microbiol 54:2735–2742. <https://doi.org/10.1128/JCM.01444-16>.
 38. Horner C, Mawer D, Wilcox M. 2012. Reduced susceptibility to chlorhexidine in staphylococci: is it increasing and does it matter? J Antimicrob Chemother 67:2547–2559. <https://doi.org/10.1093/jac/dks284>.
 39. Russell AD. 2000. Do biocides select for antibiotic resistance? J Pharm Pharmacol 52:227–233. <https://doi.org/10.1211/0022357001773742>.
 40. Maillard JY. 2007. Bacterial resistance to biocides in the healthcare environment: should it be of genuine concern? J Hosp Infect 65:60–72. [https://doi.org/10.1016/S0195-6701\(07\)60018-8](https://doi.org/10.1016/S0195-6701(07)60018-8).
 41. Williamson DA, Carter GP, Howden BP. 2017. Current and emerging topical antibacterials and antiseptics: agents, action, and resistance patterns. Clin Microbiol Rev 30:827–860. <https://doi.org/10.1128/CMR.00112-16>.
 42. Reference deleted.
 43. Kwong SM, Ramsay JP, Jensen SO, Firth N. 2017. Replication of staphylococcal resistance plasmids. Front Microbiol 8:2279. <https://doi.org/10.3389/fmicb.2017.02279>.
 44. Jensen SO, Apisiridej S, Kwong SM, Yang YH, Skurray RA, Firth N. 2010. Analysis of the prototypical *Staphylococcus aureus* multiresistance plasmid pSK1. Plasmid 64:135–142. <https://doi.org/10.1016/j.plasmid.2010.06.001>.
 45. Kwong SM, Lim R, Lebard RJ, Skurray RA, Firth N. 2008. Analysis of the pSK1 replicon, a prototype from the staphylococcal multiresistance plasmid family. Microbiology 154:3084–3094. <https://doi.org/10.1099/mic.0.2008/017418-0>.
 46. Hultner N, Ilhan J, Wein T, Kadibalban AS, Hammerschmidt K, Dagan T. 2017. An evolutionary perspective on plasmid lifestyle modes. Curr Opin Microbiol 38:74–80. <https://doi.org/10.1016/j.mib.2017.05.001>.
 47. Ramirez MS, Tolmasky ME. 2010. Aminoglycoside modifying enzymes. Drug Resist Updat 13:151–171. <https://doi.org/10.1016/j.drug.2010.08.003>.
 48. Gillespie MT, Lyon BR, Messerotti LJ, Skurray RA. 1987. Chromosome- and plasmid-mediated gentamicin resistance in *Staphylococcus aureus* encoded by Tn4001. J Med Microbiol 24:139–144. <https://doi.org/10.1099/00222615-24-2-139>.
 49. Mahairas GG, Lyon BR, Skurray RA, Pattee PA. 1989. Genetic analysis of *Staphylococcus aureus* with Tn4001. J Bacteriol 171:3968–3972. <https://doi.org/10.1128/jb.171.7.3968-3972.1989>.
 50. Rouch DA, Byrne ME, Kong YC, Skurray RA. 1987. The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and nucleotide sequence analysis. J Gen Microbiol 133:3039–3052. <https://doi.org/10.1099/00221287-133-11-3039>.
 51. McKenzie J, Hoshino T, Tanaka T, Sueoka N. 1986. The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. Plasmid 15:93–103. [https://doi.org/10.1016/0147-619X\(86\)90046-6](https://doi.org/10.1016/0147-619X(86)90046-6).
 52. Santanam P, Kayser FH. 1978. Purification and characterization of an aminoglycoside inactivating enzyme from *Staphylococcus epidermidis* FK109 that nucleotidylates the 4'- and 4"-hydroxyl groups of the aminoglycoside antibiotics. J Antibiot (Tokyo) 31:343–351. <https://doi.org/10.7164/antibiotics.31.343>.
 53. Trieu-Cuot P, Courvalin P. 1983. Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5"-aminoglycoside phosphotransferase type III. Gene 23:331–341. [https://doi.org/10.1016/0378-1119\(83\)90022-7](https://doi.org/10.1016/0378-1119(83)90022-7).
 54. Grayson ML, Russo PL, Cruickshank M, Bear JL, Gee CA, Hughes CF, Johnson PD, McCann R, McMillan AJ, Mitchell BG, Selvey CE, Smith RE, Wilkinson I. 2011. Outcomes from the first 2 years of the Australian National Hand Hygiene Initiative. Med J Aust 195:615–619. <https://doi.org/10.5694/mja11.10747>.
 55. Johnson PD, Martin R, Burrell LJ, Grabsch EA, Kirsas SW, O'Keeffe J, Mayall BC, Edmonds D, Barr W, Bolger C, Naidoo H, Grayson ML. 2005. Efficacy of an alcohol/chlorhexidine hand hygiene program in a hospital with high rates of nosocomial methicillin-resistant *Staphylococcus aureus* (MRSA) infection. Med J Aust 183:509–514.
 56. Aso Y, Sashihara T, Nagao J, Kanemasa Y, Koga H, Hashimoto T, Higuchi T, Adachi A, Nomiyama H, Ishizaki A, Nakayama J, Sonomoto K. 2004. Characterization of a gene cluster of *Staphylococcus warneri* ISK-1 encoding the biosynthesis of and immunity to the lantibiotic, nukacin ISK-1. Biosci Biotechnol Biochem 68:1663–1671. <https://doi.org/10.1271/bbb.68.1663>.
 57. Howden BP, Johnson PDR, Ward PB, Stinear TP, Davies JK. 2006. Isolates with low-level vancomycin resistance associated with persistent methicillin-resistant *Staphylococcus aureus* bacteremia. Antimicrob Agents Chemother 50:3039–3047. <https://doi.org/10.1128/AAC.00422-06>.
 58. Howden BP, Seemann T, Harrison PF, McEvoy CR, Stanton J-AL, Rand CJ, Mason CW, Jensen SO, Firth N, Davies JK, Johnson PDR, Stinear TP. 2010. Complete genome sequence of *Staphylococcus aureus* strain JKD6008, an ST239 clone of methicillin-resistant *Staphylococcus aureus* with intermediate-level vancomycin resistance. J Bacteriol 192:5848–5849. <https://doi.org/10.1128/JB.00951-10>.
 59. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
 60. Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10:e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>.
 61. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new

- genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
62. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
63. Carver TJ, Rutherford KM, Berriman M, Rajandream M, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. *Bioinformatics* 21:3422–3423. <https://doi.org/10.1093/bioinformatics/bti553>.
64. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
65. Clinical and Laboratory Standards Institute. 2016. Performance standards for antimicrobial susceptibility testing, 26th ed. Document M100S. Clinical and Laboratory Standards Institute, Wayne, PA.

4.2.2 Supplementary Material

Baines, S. L., S. O. Jensen, N. Firth, A. Gonçalves da Silva, T. Seemann, G. P. Carter, D. A. Williamson, B. P. Howden and T. P. Stinear (2019). "Remodelling of pSK1 family plasmids and enhanced chlorhexidine tolerance in a dominant hospital lineage of methicillin-resistant *Staphylococcus aureus*." *Antimicrobial Agents and Chemotherapy*; DOI: 10.1128/AAC.02356-18.

Available in Appendix 1: Supplementary Dataset. This file summarises the relevant isolates demographics and WGS information including sequence data accession numbers for all isolates included in this study.

Supplementary Results: Part A. Description of the pSK1-like plasmid variants

A description of the pSK1-like plasmid variants, as identified through long-read sequencing, has been described in the main text and FIG 2. Provided here is further detail about these plasmid structures and how the insertion sequence (IS) features, specifically the non-coding regions (NCR), the terminal inverted repeats (TIR), the target site duplications (TSD), and their location relative to plasmid and chromosomal genes were used to identify which pSK1-like SV was likely carried by each ST239 isolate.

SV1: Absence of Tn4001. In pSK1, Tn4001 is flanked by inverted copies of IS256 with perfect 8 bp TSD (TAAGTAAA). In SV1 only a single copy of this TSD sequence was identified in isolation from an IS256 NCR at the expected plasmid location for Tn4001 (FIG 2A). Tn4001 is highly mobile and has been found at multiple chromosomal and plasmid locations (375, 376). In BPH2056 (SV1 representative genome), a chromosomal copy of Tn4001 was identified, flanked by perfect 8 bp TSD (CAAAATCA), located downstream of a hypothetical protein in close proximity to the DNA-directed DNA polymerase III alpha subunit (*dnaE*) gene. Using this variation in TSD patterns flanking Tn4001, it was identified that of the ten short-read assemblies which demonstrated an orthologue pattern consistent with SV1 (FIG S1), five likely carried a plasmid consistent with pSK1 (with Tn4001 located in the plasmid), and the other five likely carried a plasmid consistent with SV1, with a chromosomal copy of Tn4001 identified in four isolates.

SV2: IS256-Mediated Chromosomal Integration & Loss of 5,175 bp. Chromosomal integration of SV2 was mediated by IS256, with two inverted copies flanking the 24,299 bp region, termed IS256-R and IS256-2. A third copy (IS256-L) is present as part of Tn4001 (FIG 2B). Additionally, SV2 has a 5,175 bp deletion with loss of six genes when compared to pSK1 (FIG 2B). The upstream TSD for IS256-2 and IS256-R (external to SV2) were identical and represented a novel sequence (GAGTTGAC). A second novel TSD was identified downstream of IS256-2 (internal to SV2, ATGATTTC), with a single copy of the same sequence identified in pSK1 (not adjacent to an IS256 NCR) at the boundary of the region absent in SV2 (FIG 2B). This TSD pattern was identified in all seven short-read assemblies which demonstrated an orthologue pattern consistent with SV2, suggesting that they all carried this integrated plasmid variant.

SV3: IS257-Mediated Deletion of Tn4003 & IS256-Mediated Inversion. In pSK1, Tn4003 is flanked by direct copies of IS257 with perfect 8 bp TSD (TTTTATAA). In SV3, only a single copy of IS257 was identified, flanked

by the same TSD (FIG 2A). Additionally, SV3 also had a 3,488 bp inversion compared to pSK1, flanked by IS256-R and IS256-3 (FIG 2A). The upstream TSD for IS256-R (internal to the inversion) and the downstream TSD for IS256-3 (external to the inversion) represented a novel sequence (TTTGAGTT), the former was identified in the reverse complement. A single copy of the same sequence (TTTGAGTT) was identified at the boundary for the inverted segment in pSK1 (FIG 2A, indicated by the dark red circle). The upstream TSD for IS256-3 (internal to the inversion, TTTACTTA) was the same sequence as the upstream TSD for IS256-R in pSK1 (TAAGTAAA), but in the reverse complement (FIG 2A). Utilising the variation in TSD pattern, including the change in sequence orientation and adjacent plasmids genes, it was identified that only seven of the eleven short-read assemblies which demonstrated an orthologue pattern consistent with SV3 were likely to carry this plasmid variant.

SV4: IS256-Mediated Chromosomal Integration & Loss of 3,488 bp. Chromosomal integration of SV4 was mediated by IS256, with two inverted copies flanking the 22,047 bp region. A 3,488 bp region, specifically the region inverted in SV3 was absent (FIG 2B). All internal TSDs were found to correspond with those identified in SV3. The TSD upstream of IS256-R and IS256-3 (external to SV4) were novel (CTTTGTAT) (FIG 2B). This TSD pattern was identified in all four short-read assemblies which demonstrated an orthologue pattern consistent with SV4.

SV5': IS257-Mediated Chromosomal Integration & Loss of 404 bp, SV5: IS257-Mediated Inversion & Disruption of Replication Machinery, & SV6: IS256-Mediated Deletion of 5,105 bp. A hypothesised structure has been developed for SV5', which represents an intermediate structure in the development of SV5 and SV6 (FIG 2). Chromosomal integration in SV5' would likely be mediated by IS257, inserting 9,233 bp upstream of a disrupted copy of the β -haemolysin gene. For compatibility with the recovered SV5 structure, this region would be flanked by IS257-R1 and IS257-R2 from pSK1, located in the same orientation, with absence of the intervening 404 bp region (FIG 2C). All TSD that are internal to SV5' would be unaltered from the sequences identified in pSK1. The TSD downstream of IS257-R2 and IS257-R1 (external to SV5') would represent a novel perfect 8 bp TSD (GTTTTTAC).

In SV5, the intermediate SV5' has undergone a 42,660 bp inversion (FIG 2C). This would likely be the result of intramolecular replicative transposition in the inverse orientation (54), and result in the creation of a novel pair of 8 bp TSD (TTCTTATC), located adjacent (in opposite orientations) to both copies of IS257-R2*, one external to the inverted region (adjacent to the unaffected region of ϕ Sa3) and the other one internal (adjacent to the inverted region of ϕ Sa3) (FIG 2C). In addition to the inversion, a 45 bp deletion in *repA* was detected in SV5, causing fragmentation of the gene annotation.

In a separate event, the intermediated SV5' structure has undergone an IS256-mediated 5,105 bp deletion event, resulting in the 22,520 bp structure SV6 (FIG 2B). All TSD are identical to those presented in the hypothesised SV5' except for one; the upstream TSD for IS256-R (AAATTGTG) is novel. A single copy of this TSD was identified in pSK1, not adjacent to an IS256 NCR at the boundary for the deletion (FIG 2A).

Of the remaining 63 short-read assemblies, four had an orthologue pattern consistent with SV3, 21 with SV5, and 38 with SV6. Using the variation in TSD patterns described above, it was identified that 24 isolates likely harboured an integrated plasmid consistent with SV5', three with the inverted SV5, and 36 with the shortened

SV6. All 63 isolates formed a single monophyletic clade from the expected ancestral node in which SV5' is predicted to have emerged (FIG 3). Further, these classifications were completely congruent with the phylogenetic tree, consistent with the proposed development of these plasmid variants (FIG 3). All isolates harbouring an SV5'-like or SV5-like plasmid variant appeared to have disrupted plasmid replication machinery, the large majority demonstrating a fragmented *repA* gene annotation (FIG S1). Two isolates had more extensive deletions, with loss of multiple genes similar to SV6 (FIG 3 & FIG S1). These two isolates appear to represent two further independent IS256-mediated deletion events that have occurred following chromosomal integration. Similarly, four isolates appeared to have undergone an IS257-mediated deletion of *Tn4003*, with their location in the phylogeny suggestive of two independent events (FIG 3 & FIG S1).

Supplementary Results: Part B. Investigation of the association between plasmid evolution and biocide tolerance

Provided here are the additional investigations undertaken to explore the potential association between the emergence of the pSK1-like plasmid SVs and the development of chlorhexidine tolerance in the ST239 population.

Mutations in *QacAR*. A comparison of all *qacAR* sequences (n = 89 pSK1-like, n = 67 pTW20_1-like) to reference JKD6008 (Australian clade) identified six SNPs resulting in missense mutations: five in *QacA* and one in *QacR*. The most common mutation resulted in a *QacA* T₃₅₈I amino acid (aa) change. This SNP was identified in all 67 *qacA*-containing Asian-Australian clade isolates and none of the *qacA*-containing Australian clade isolates. As this mutation was invariably associated with one plasmid type, irrespective of MIC, it represented the distinct evolutionary histories of the *qacA* genes in these different plasmid populations and was unlikely to be contributing to phenotypic variation. The only other two commonly observed mutations were identified in the Australian clade and resulted in a *QacA* H₁₆₄Q or a *QacA* V₄₀₈I aa change. These mutations were phylogenetically correlated, *QacA* H₁₆₄Q was identified in 14/18 isolates that harboured SV2, SV3 or SV4, and *QacA* V₄₀₈I was identified in 8/13 isolates that harboured SV3 or SV4. The presence of these mutations did not correlate with CHX phenotype, suggesting that these SNPs were most likely markers of clonal evolution and not significant contributors to the larger shift in CHX tolerance.

Plasmid and *qacA* copy number. Intermediate sized staphylococcal plasmids like pSK1 are generally considered to be low copy number plasmids (288). Multiple copies of *qacA*, resulting from increased plasmid or gene copy number could potential be contributing to the phenotypic variation observed between the Australian and Asian-Australian clades. To investigate this, *qacA* gene copy number was estimated by comparing the mapped sequence read depth of contigs representing the plasmid regions (as defined in FIG S2) to those representing the chromosome. The ratio of mapped-reads in the plasmid contigs representing region 4 (containing *qacAR*) and region 1 (plasmid backbone) were 0.89 ($\pm$ 0.18) in the pSK1-like population and 1.12 ($\pm$ 0.41) in the pBPH2003-like (pTW20_1-like) population. This suggested that each plasmid copy contained only a single copy of *qacA*. The ratio of mapped-reads in these same contigs compared to contigs representing the chromosome (specifically those containing *dnaA*) were 0.68 ($\pm$ 0.28) and 0.77 ($\pm$ 0.24) in the pSK1-like population, and 2.81 ($\pm$ 1.32) and

2.69 ($\pm$ 1.33) in the pBPH2003-like population for region 4 and region 1 respectively. These ratios suggested that the pSK1-like plasmids were typically single copy in the Australian clade isolates, consistent with most isolates harbouring a chromosomally integrated variant. No significant differences were observed between isolates that harboured an extra-chromosomal or integrated pSK1-like variant. Conversely, the Asian-Australian isolates carried one to four copies of a pBPH2003-like plasmid. As this clade has a lower average MIC to chlorhexidine (TABLE 2), plasmid copy number did not appear to be contributing to the variation observed in phenotypic CHX susceptibility.

Discriminant Analysis of Principle Components. DAPC is a multivariate technique that can be used to reconstruct population subdivisions hypothesised to be present using genomic data; it finds the smallest number of principle components to build discriminant functions that minimise within group variation and maximise between group variation (377). This approach was utilised to identify potential associations between phenotypic CHX MIC and the presence and/or absence of specific genes. In this analysis the population subdivisions hypothesised to be present in the dataset were the CHX MIC values (1, 1.5, 2, 3, 4, and 6 mg/L). Two models were generated: (i) one dataset representing all accessory genes identified in the ST239 collection ($n = 212$), and (ii) the same dataset with the pSK1 and pBPH2003 (pTW20_1-like) genes excluded.

In the first model, broad clusters were able to be resolved that reflected the low (1, 1.5, and 2 mg/L), medium (3 mg/L) and high (4 and 6 mg/L) MIC values (FIG S4 A, B, C). The top contributing variables in this separation were *qacAR* and the co-located genes in plasmid region 4 in the first discriminant function, and the genes specific to pBPH2003 in the second discriminant function (FIG S4 D). This finding is consistent with the increase in MIC that was associated with plasmid acquisition and the distinct CHX susceptibility profiles of the pSK1-like and pBPH2003-like populations (TABLE 2). Furthermore, as all but one of the *qacA*-negative isolates belonged to the Australian clade, the second discriminant function was also reflective of the phylogenetic division in the ST239 population (FIG S4 E).

In the second model the MIC-reflective clusters were less defined (FIG S4 A, B, C). There was some separation based on chromosome-associated gene orthologues, none of which appeared to be associated with efflux systems based on gene annotations (not shown). The membership profiles of the isolates in this model (a measure of assignment probability to the prior hypothesised subdivisions) appeared to be strongly correlated with the phylogeny (FIG S4 E), which suggested that this DAPC model was detecting the underlying phylogenetic divisions within the population, which correlated with both CHX MIC and emergence of the pSK1-like plasmid variants.

Genome Wide Association Studies. A GWAS analysis was utilised to explore whether mutation in the core genome were associated with CHX MIC. This dataset was tested at three different phenotypic divisions, with the tolerant population defined as a MIC ≥ 3 , ≥ 4 or 6 mg/L. Numerous mutations were identified as significantly associated with MIC after correcting for multiple hypothesis testing (FIG S5 A). Examination of the 50 most significant SNPs for each phenotypic division found all to be strongly correlated with the phylogenetic model for the ST239 population (FIG S5 B). GWAS approaches are highly susceptible to phylogenetically correlated phenotypes and even if a phylogenetic penalty was applied, it may not be possible to detect a SNP in this dataset that is responsible for increased CHX MICs using this approach.

The efflux system encoded by *norA* can actively export CHX and mutations in this gene, and a homologue *norB*, have been associated with enhanced CHX tolerance (180, 198, 199). Mutations were detected in both *norA* and *norB* in this dataset, however none were significantly associated with CHX MIC in any phenotypic division tested.

Supplementary Materials & Methods

Bacterial Isolates. This study utilised a collection of 212 Australian ST239 *S. aureus*, selected from available collections for temporal and geographic diversity. Isolates were recovered between 1980 and 2012 from the Australian states of Victoria (n = 123), New South Wales (n = 62), Queensland (n = 22), South Australia (n = 2), and Western Australia (n = 1). Two isolates, including reference *S. aureus* JKD6008, were recovered in New Zealand (305). Bacterial isolates and corresponding whole genome sequence (WGS) data were sourced from:

- The complete genome of JKD6008 was sourced from (325).
- 73 isolates were sourced from (137). Six isolates from the original study were excluded because they represented secondary samples from patients already represented in the collection. These isolates have only been included in the generation of the global phylogenetic model (described below).
- 104 isolates were sourced from the collection of the Australian and New Zealand Cooperative on Outcome in Staphylococcal Sepsis (ANZCOSS) study (378). All ST239 from unique patients selected for the ANZCOSS sub-study were included (306).
- 16 isolates were sourced from the collection of the Vancomycin Efficacy in Staphylococcal Sepsis in Australasia (VANESSA) study (336). All ST239 isolates from unique patients were included.
- 18 isolates were sourced from the historic collections (1989-1994) housed at the Microbiological Diagnostic Unit Public Health Laboratory, Melbourne, Australia. All ST239 isolates from unique patients in which a viable bacterial isolate could be recovered were included.

To generate a global phylogenetic model for the ST239 lineage the Australian ST239 collection was supplemented with the publicly available WGS data for 319 globally diverse ST239 isolates, selected to maximise temporal, geographic and genetic diversity of the collection. The WGS data for these isolates was sourced from:

- The complete genomes of eight ST239 isolates were sourced from (216, 322-324, 326-329).
- 174 isolates were sourced from the collections of (300, 301, 330-334). All ST239 isolates that could be accessed from these studies were included.
- 78 isolates were sourced from the collections of (31, 303). A limited number of isolates were included from both studies, selected to maximise the temporal and geographic diversity of these collections
- 29 isolates were sourced from the 2015 Australian Staphylococcal Sepsis Outcome Program (ASSOP). Sequence data for all ST239 isolates was kindly provided by the Australian Group on Antimicrobial Resistance (<http://www.agargroup.org/>).
- A further 24 isolates were sourced from historic collections (1989 – 1994) housed at the Microbiological Diagnostic Unit Public Health Laboratory, Melbourne, Australia.

Whole Genome Sequencing & Sequence Data. The WGS data for 42 isolates is novel to this study (Supplementary Dataset). Genomic DNA was extracted with the JANUS Automated Workstation, using the Chemagic DNA/RNA kit (Perkin Elmer). Unique dual index libraries were prepared using the Nextera XT DNA preparation kit (Illumina). Libraries were sequenced on a NextSeq platform (Illumina) with 2 x 150 bp chemistry. Long read sequencing was conducted on seven isolates (BPH2003 [pTW20_1-like], BPH2019 [SV3], BPH2056 [SV1], BPH2070 [SV2], BPH2869 [SV6], BPH3244 [SV4], and JKD6009 [SV5]). Genomic DNA was extracted using the DNeasy Blood & Tissue extraction kit (Qiagen) and sequencing was conducted on the Pacific Biosciences RS-II platform with P6-C4 chemistry. WGS data accession numbers can be found in the Supplementary Dataset.

Genome Assembly. Long-read sequence data was *de novo* assembled in the SMRT Analysis System v2.3.0.140936 (Pacific Biosciences), using the protocol previously described (194). Briefly, raw sequence data was first assembled using the HGAP v3 protocol, with a minimum seed read length of 10,000 bp, genome size of 3 Mb, target coverage of 10 and an overlapper error rate of 0.04. Contigs underwent a second round of error correction using Quiver v1, then were circularised with the overlap removed and orientated to *dnaA* or *repA* in Geneious v8.1.5 (Biomatters). Long reads were aligned to the assembly using BridgeMapper v1 and manually checked for assembly errors. The consensus sequence from this alignment was further error corrected with high quality short-read Illumina data using Snippy v3.2 (<https://github.com/tseemann/snippy/>), performed in an iterative process until no variants were detected. The final sequences were annotated with Prokka v1.12. Structural comparison of the complete plasmid sequences were conducted using the Artemis Comparison Tool (379), filtering out hits < 100 bp in length.

Short read sequence data was *de novo* assembled using SPAdes v3.11.0 (380), excluding the pre-assembly error correction (“--only-assembler”), including post-assembly error correction (“--careful”) and a minimum coverage threshold of five (“--cov-cutoff”) to remove potential contaminants. Assemblies were annotated using Prokka v1.12 (337) and contigs < 200 bp in length were excluded. Contig coverage information was extracted from the assembly data, with the ratio of coverage between chromosome and plasmid associated contigs used to approximate plasmid and gene copy number.

Orthologue Clustering & Gene Synteny. The presence/absence of plasmid gene orthologues was used to screen short-read assemblies for the presence of a pSK1-like plasmids. The pangenome pipeline Roary (381) was used to perform the ortholog clustering. The protein coding sequences (CDS) identified in the annotated complete genomes and short-read *de novo* assemblies were clustered into orthologous groups at different thresholds of amino acid homology (75%, 80%, 85%, 90%, and 95%), without splitting paralogs (option “-s”). No significant variation was observed in the clustering across the different homology thresholds (data not shown), subsequently all analyses published here use a 95% amino acid homology threshold. Genes were deemed to be syntenic if the CDS were identified as adjacent on a single contig.

Phylogenetic Analysis. The Illumina short-read sequence data for the 211 Australian ST239 *S. aureus* were mapped to the complete genome of reference JKD6008 [CP002120, (325)], using Snippy v3.2 (<https://github.com/tseemann/snippy/>). An alignment of core genome SNPs was generated using a minimum coverage of 10 reads, and base-call stringency of 90%. SNPs were deemed core if a base was called for every

isolate in the alignment at that site. This alignment was used as the input to generate a ML phylogenetic tree with IQ-TREE v1.6.1 (339), modelled with a GTR nucleotide substitution model, empirical base frequencies, a correction for the absence of invariant sites, use of a discrete Gamma model with four rate categories (model option “GTR+F+I+G4”), and 1,000 ultrafast bootstrap replicates (option “-bb 1000”).

Two additional mapped alignments were generated: (i) to represent the global ST239 collection (n = 531, mapped to JKD6008), and (ii) to represent the Australian ST239 clade (n = 124, mapped to JKD6008) using the process described above. Recombination was assessed using ClonalFrameML v.1.11 (338). However, as very few recombinant events were detected (data not shown) masking these sites prior to temporal phylogenetic analysis was deemed unnecessary. The package BEAST v2.4.7 (340) was used to construct the time-aware phylogenetic models. From the whole genome alignments, the core genome SNPs were extracted, grouped into a single partition and modelled with the following parameters. The year of isolation was provided as a prior. The GTR model of nucleotide substitution was selected, with the substitution rate estimated using four gamma categories. A relaxed lognormal clock model and a coalescent constant population model were selected (based on previous ST239 studies (137, 299)). Invariant sites were included as a constant pattern, weighted based on the nucleotide proportions of reference JKD6008. For the Australian clade, chlorhexidine MIC data was modelled as a discrete trait. Five 50 million step iterations were run and examined for convergence in Tracer v1.6 (<http://tree.bio.ed.ac.uk/software/tracer/>). The burn-in was adjusted for each run to a point at which a stable likelihood was achieved. All runs converged, and none were discarded, based on the comparisons of Bayes Factors, calculated from the harmonic means of the run likelihoods, and the attainment of effective sample sizes (> 200) in all measurements. Runs were joined and re-sampled at a frequency that would achieve ~10,000 trees using LogCombiner. Maximum clade credibility trees with median node heights were generated in TreeAnnotator.

Genome Wide Association Study. Allelic association was performed using the GWAS tool Plink (317). Core genomes SNPs (identified as described above) were used as the input. Three phenotypic divisions were tested, with the tolerant populations defined as a MIC ≥ 3 , ≥ 4 or 6 mg/L. A Bonferroni correction was used to account for multiple hypothesis testing, the significance threshold determined as 0.05/number of SNPs tested.

Discriminant Analysis of Principal Components. DAPC was implemented in R v3.4.2 (<http://www.R-project.org/>) using the package adegenet (382). Presence or absence of the CDS ortholog groups (identified as described above) were used as the input for the DAPC models, filtering out core ortholog groups (present in ≥ 210 isolates) and singletons groups (present in ≤ 3 isolates). Two models were generated: (i) all CDS ortholog groups that met the input criteria, and (ii) the same data set excluding all plasmid associated CDS. In both models, isolates were assigned to prior populations based on chlorhexidine MIC values (1, 1.5, 2, 3, 4, and 6 mg/L). DAPC models were assessed individually to determine the optimum number of principle components to be retained, based on the a-score and cross validation functions in adegenet.

Antimicrobial Susceptibility Testing. Phenotypic susceptibility testing for gentamicin and trimethoprim was performed using E-tests (bioMérieux), following the manufacturer’s recommendations. *S. aureus* isolate ATCC29213 was used as a control. All results were read in triplicate and interpreted using CLSI guidelines (M100S, 26th Ed).

Phenotypic susceptibility testing to CHX was performed by broth microdilution, as per the CLSI protocol for antimicrobial susceptibility testing (M07A10E, January 2015). Chlorhexidine digluconate was sourced from Sigma-Aldrich (#C9394) and diluted in sterile H₂O to 128 mg/L. In microtiter trays (Costar, #3799), serial dilutions of 100µl of CHX in cation adjusted Muller Hinton (CaMH) broth (Oxoid) were generated to a final concentration range of 1.0 - 32 mg/L, tested in 1.5x increments. A second concentration range of 0.25 - 8 mg/L was also used to test isolates found to have MIC values < 2 mg/L. Bacterial isolates, from a fresh overnight plate culture, were resuspended in saline and adjusted to a starting concentration equivalent to a 0.5 McF solution, then diluted 1:100 in CaMH broth. In a 1:1 ratio, 100µl of the bacterial suspension was added to each well, and 200µl to an empty growth control well. Microtiter plates were incubated for 24 hours in a static condition at 37°C. The MIC was interpreted as the well of lowest concentration with no visible growth.

To assess viability post CHX exposure (approximate MBC determination), after the MIC was read 10µl from each well in the microtiter plate was sub-cultured to a CaMH agar plate (Oxoid). No CHX inhibitors were added. Plates were incubated for 24 hours at 37°C. The MBC was interpreted as the lowest concentration in which no growth was observed, growth being defined as ≥ 2 colonies. All isolates were tested in biological triplicate. A fourth BMD was performed if the first three replicates were found to have a range in excess of 2 wells in the MIC, or 4 wells in the MBC. The median values were used for statistical analysis. Controls were tested regularly and a minimum of twice for every batch performed from a fresh dilution of CHX. Controls included *S. aureus* isolate ATCC29213, JKD6008 and culture-negative wells.

All statistical analyses were performed in R v3.4.2 (<http://www.R-project.org/>). Prior to significance testing all MIC and MBC values were transformed to a log₂ scale. Comparison of MIC or MBC values between two population were conducted using a Welsh two sample t-test. Simple linear models were generated to assess temporal trends, comparing year of isolation against log₂ transformed MIC and MBC data using the function “lm”. Significance was determined as a p value ≤ 0.05 .

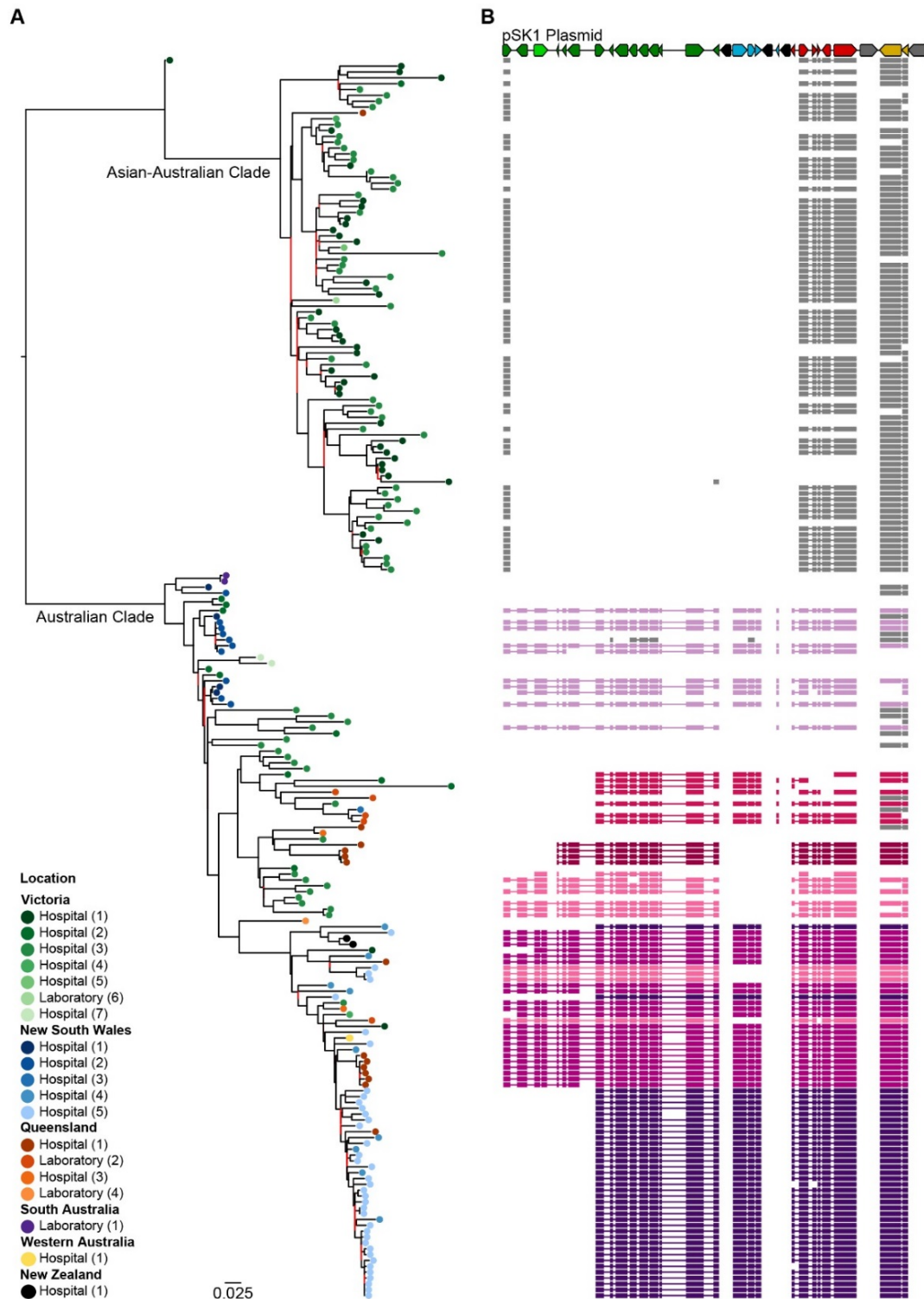


FIG S1 pSK1 plasmid gene presence and synteny. (A) Maximum likelihood phylogenetic tree inferred from 3,883 core genome SNPs illustrates the population structure of ST239 *S. aureus* in Australia. Tips are coloured based on location (refer to key). Branches with < 70% bootstrap support are coloured red. (B) Coloured blocks represent the identification of a pSK1 gene orthologue, using a 95% amino acid identity threshold (excluding insertion sequences). Box length is reflective of gene length and ordered based on pSK1 (FIG S2 A). Boxes are linked if orthologues were found to be syntenic. Coloured boxes reflect the six identified patterns.

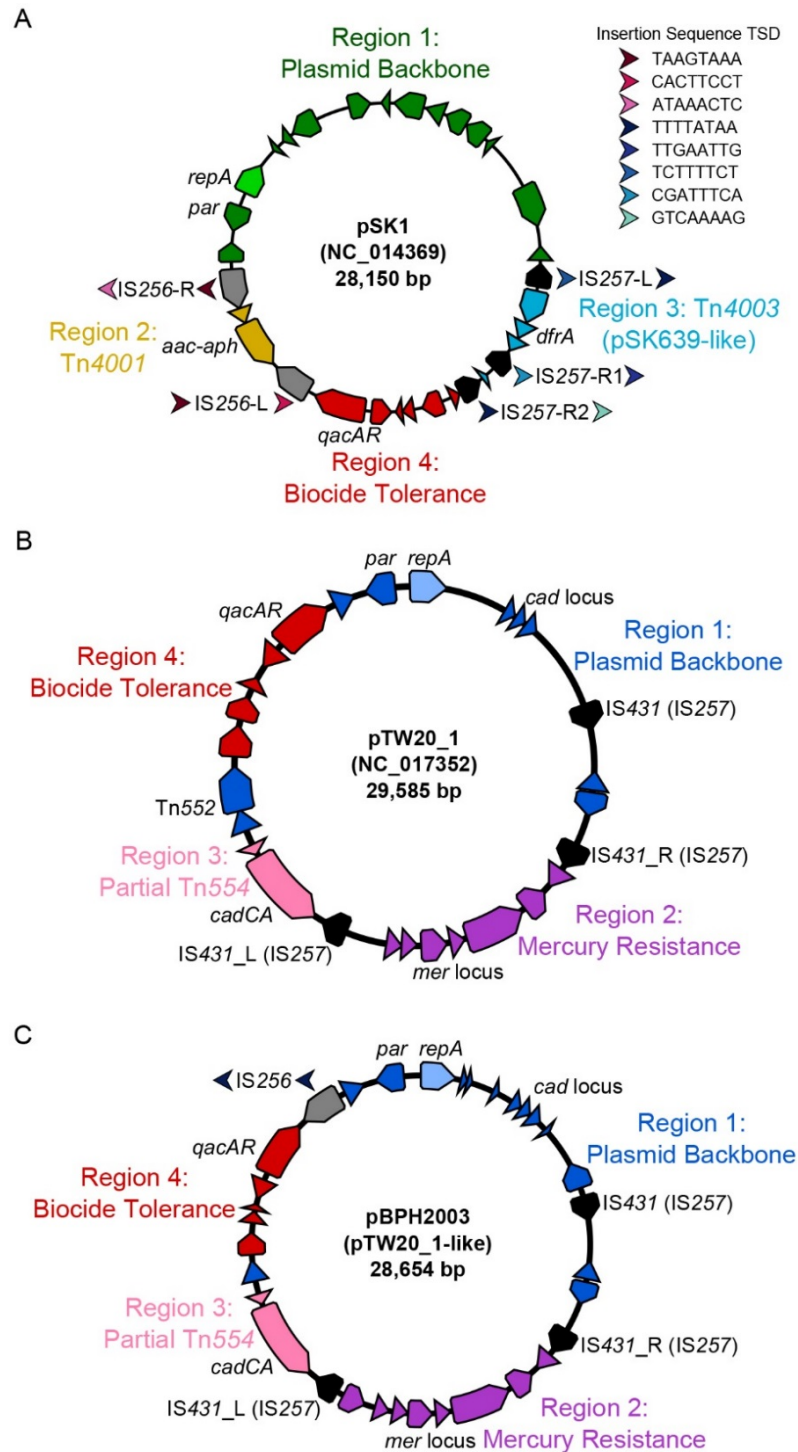


FIG S2 Structure of pSK1, pTW20_1 and pBPH2003 plasmids. Plasmid genes have been coloured based on defined plasmid regions. Insertion sequences (IS) are coloured grey (IS256) and black (IS257), with target site duplications (TSD) illustrated: arrows indicate upstream/downstream sequences, orientation, and are coloured to represent unique sequences (refer to key). (A) pSK1; sequence and annotations are that previously published, accession NC_014369 (52). (B) pTW20_1; sequence and annotations are that previously published, accession NC_017352 (29). (C) pBPH2003; a pTW20_1-like plasmid representative of that harboured in isolates from the Asian-Australian ST239 clade.

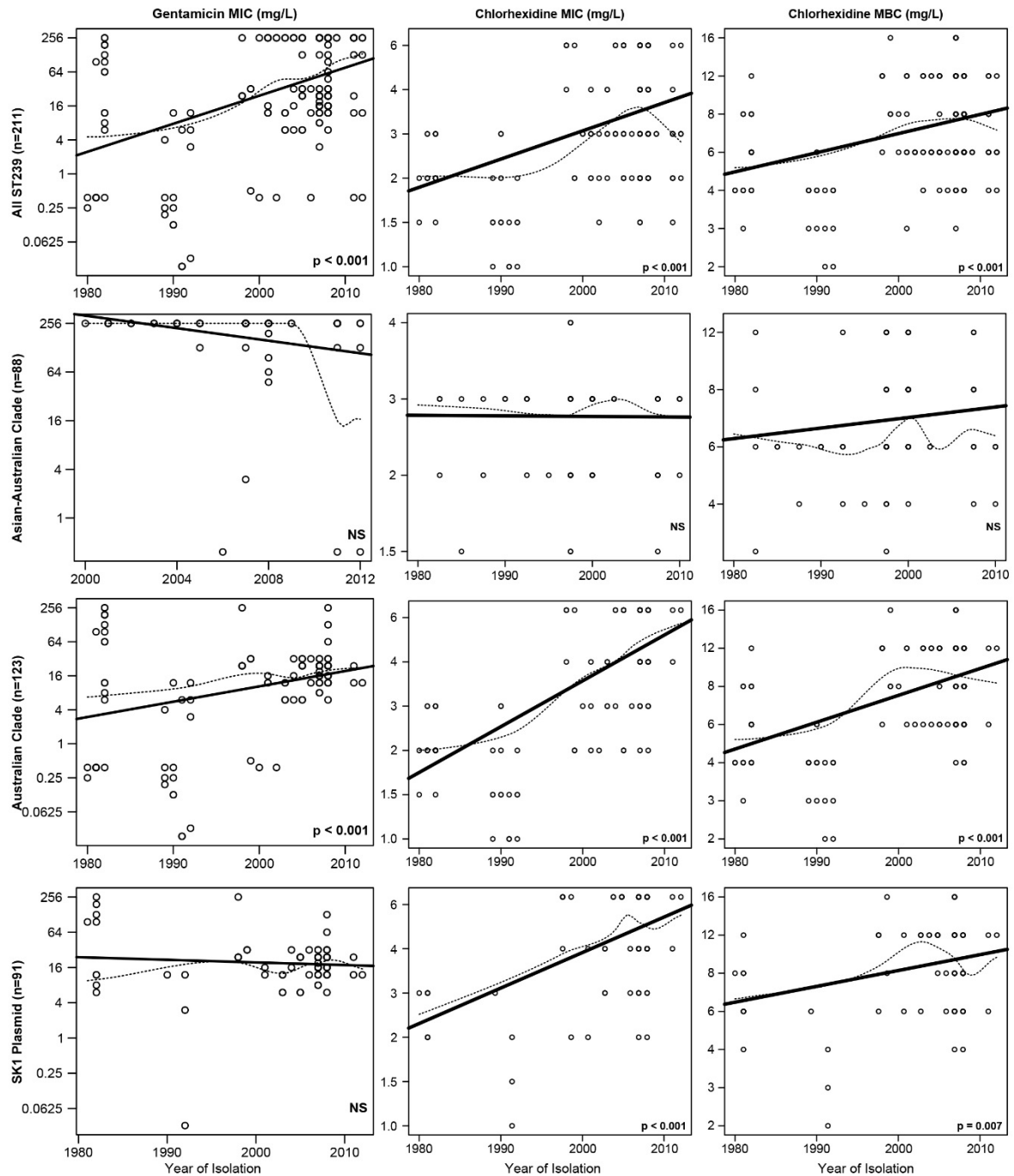


FIG S3 Modelling temporal-association in phenotypic susceptibility data. Graphs depict linear models developed to explore the potential association between gentamicin MIC and chlorhexidine MIC and MBC with the year in which isolates were recovered. The dotted line indicates the smoothed mean MIC and the bold line indicates the fitted linear model. Four populations were tested (from top to bottom): (i) All ST239 MRSA ($n = 211$), (ii) the Asian-Australian clade ($n = 88$), (iii) the Australian clade ($n = 123$), and (iv) the pSK1-like plasmid harbouring population ($n = 91$). Note: reference JKD6008 (Australian clade) was not tested.

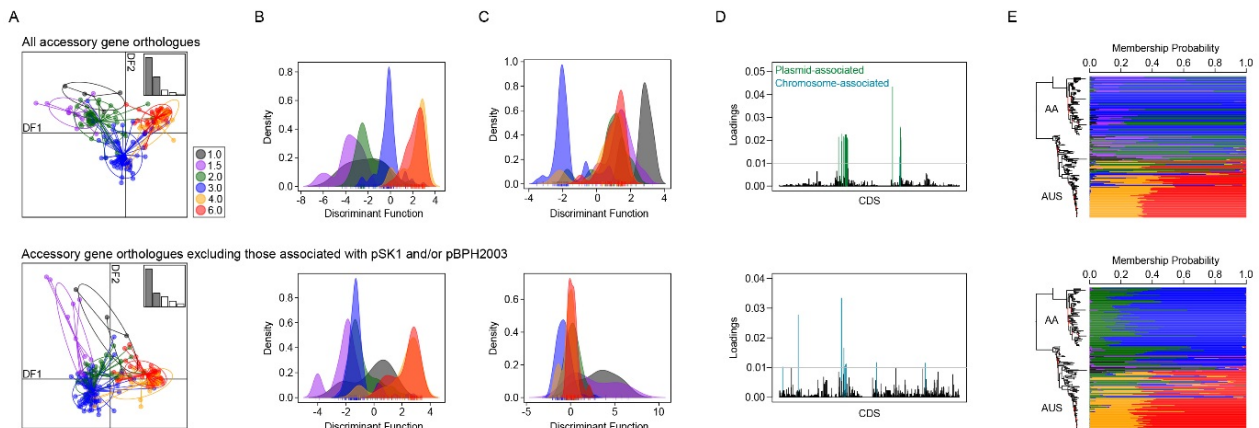


FIG S4 Exploration of gene presence/absence signatures associated with phenotypic chlorhexidine tolerance. The graphs illustrate the findings from two Discriminant Analysis of Principle Components models, used to investigate genetic signatures in the ST239 MRSA population associated with chlorhexidine MIC. The first model (top panel) examined all accessory gene orthologues (clustered at 95% amino acid homology), and in the second model (bottom panel) gene orthologues associated with either pSK1 or pBPH2003 (representative of the pTW20_1-like plasmid recovered from the Asian-Australian clade) were excluded. The prior hypothesised sub-divisions represent the phenotypic CHX MIC values for the population, as indicated in the key. (A) Scatter plots illustrate the clustering of isolates across the two most discriminant functions. The density plots illustrate the same data in a one-dimensional format for the first (B) and second (C) most discriminant functions. (D) Loading plots illustrate the contributions of all gene orthologues to the DAPC models, with the most highly contributing variables (> 0.01) coloured based on whether they represent a plasmid-associated (green) or chromosomal gene orthologue (blue). (E) Membership graphs illustrate the membership profiles for all isolates, aligned to a maximum likelihood phylogenetic tree for the ST239 population. The probability index indicates the likelihood of assigning an isolate back to one of the hypothesised sub-divisions.

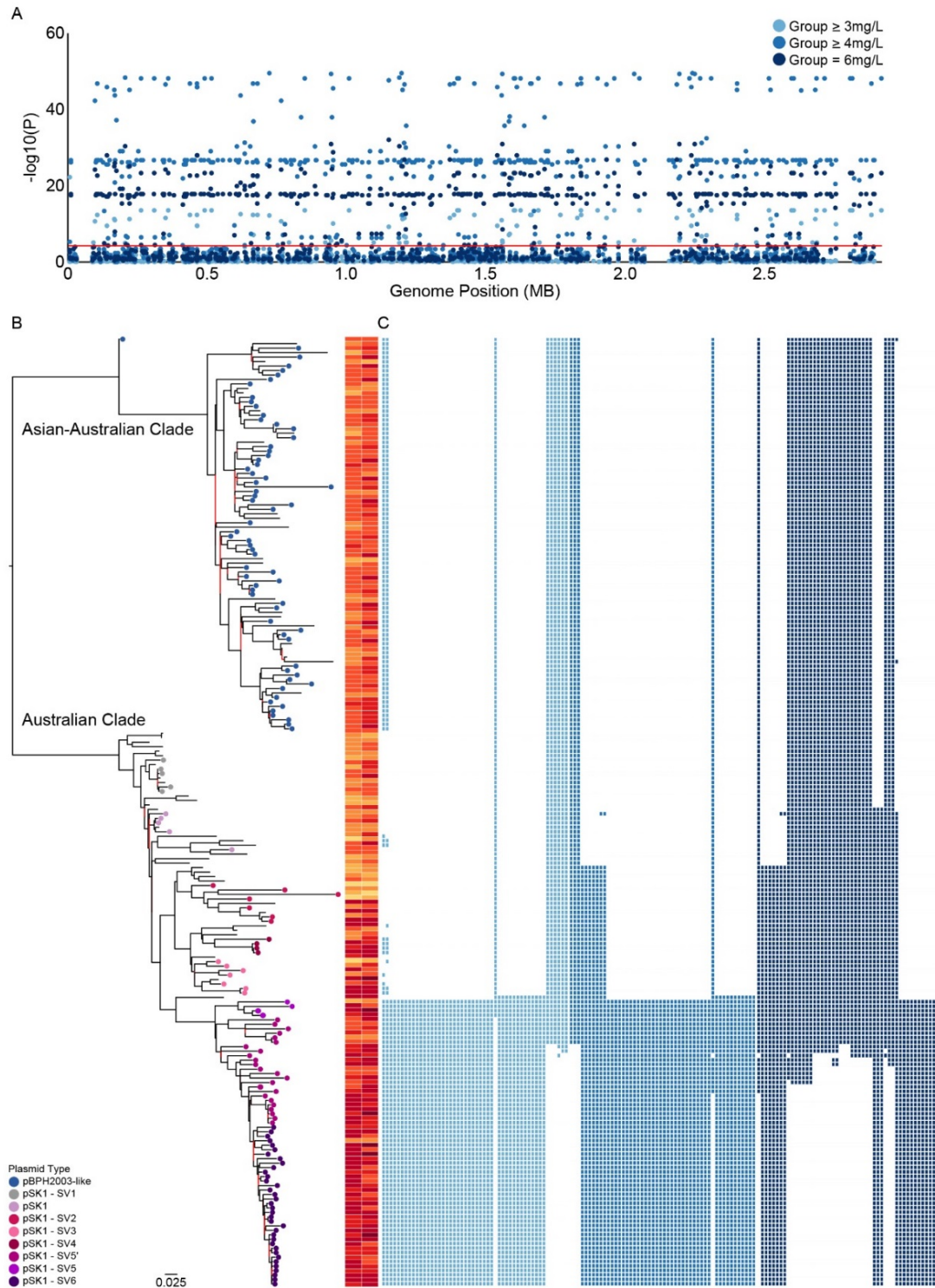


FIG S5 Exploration of mutations associated with phenotypic chlorhexidine tolerance. (A) Manhattan plot illustrates the results of three Genome Wide Association Studies conducted to identify core genome SNPs associated with CHX MIC, tested at three thresholds: MIC ≥ 3 mg/L, ≥ 4 mg/L or = 6 mg/L. Red line indicates the threshold for significance (with Bonferroni correction). (B) Maximum likelihood phylogenetic tree is that from FIG S1, and adjacent are the phenotypic CHX MIC and MBC values (coloured yellow to red with increasing value). (C) The presence (blue) or absence (white) of the 50 most significantly associated mutations detected at each phenotypic threshold are illustrated (refer to key).

TABLE S1. Investigation of gentamicin resistance in Australian ST239 MRSA

Median (mg/L)			Median (mg/L)		
Isolate Populations (n)	MIC		Isolate Populations (n)	MIC	P value
All ST239 (211)					
Australian Clade (123)	10.30	vs	Asian-Australian Clade (88)	179.80	< 0.0001
No AME (91)	0.38	vs	aac(6′)-aph(2″) (181)	58.20	< 0.0001
		vs	aadD (73)	15.20	< 0.0001
		vs	aph(3′)-III (95)	190.50	< 0.0001
Australian Clade (123)					
No AME (19)	0.39	vs	aac(6′)-aph(2″) (99)	21.90	< 0.0001
		vs	aadD (73)	15.20	< 0.0001
		vs	aph(3′)-III (8)	164.90	< 0.0001
Asian-Australian Clade (88)					
No AME (1) ^a	0.38	vs	aac(6′)-aph(2″) (82)	189.70	-
		vs	aph(3′)-III (87)	193.00	-

Abbreviations: AME, aminoglycoside modifying enzyme; MIC, minimum inhibitory concentration.

^a Statistical analysis was not performed as only a single isolate was identified harbouring no AME genes.

TABLE S2. Gentamicin resistance and chlorhexidine tolerance profiles

pSK1 Plasmids (n)	Gentamicin MIC	Median (mg/L)	
		Chlorhexidine MIC	Chlorhexidine MBC
Structural Variants			
No pSK1 Plasmid (32)	1.80	1.80	4.60
SV1 (5)	142.10*	2.80	7.30
pSK1 (5)	9.60	2.40	6.20
SV2 (7)	9.20	3.30	7.30
SV3 (7)	15.90	4.40	9.70
SV4 (4)	15.30	4.90	10.80
SV5' (3)	15.10	5.20	10.50
SV5 (24)	16.00	4.00	8.90
SV6 (36)	21.50	5.00	8.80
Structural Changes Grouped			
Extra-chromosomal (17) ^a	26.10	3.20	7.80
Integrated + <i>repA</i> deletion (25) ^b	19.10	3.70	8.50
Integrated + MG deletion (49) ^c	18.60	4.70	8.60

Abbreviations: MIC, minimum inhibitory concentration; MG, multi-gene; SV, pSK1-like structural variant.

^a Includes all isolates classified as harbouring SV1, prototypical pSK1 and SV3.

^b Includes all isolates classified as SV5' and SV5 with a fragmented pSK1 *repA* sequence identified.

^c Includes all isolates classified as SV2, SV4, SV5', SV5 and SV6 with a multi-gene deletion in pSK1 region 1.

* All five isolates representing SV1 carry *aph(3')-III*.

4.3 Summary

Mobile genetic elements have the capacity to cause significant adaptive change in bacterial populations, however the way in which they interact with and influence these populations can be subtle. This is clearly demonstrated in the work presented in this chapter. The acquisition of pSK1 early in the evolution of the Australian ST239 MRSA clade, a discernible adaptive shift, introduced multiple mechanisms of antimicrobial resistance into the ST239 population, providing distinct changes in phenotypic susceptibility to gentamicin and chlorhexidine. However, more subtle changes in the configuration and location of these plasmids appears to be the cause, or at least contributing to further enhancement of chlorhexidine tolerance. Long-read sequencing and thorough reconstruction of plasmid structures and IS elements was key in revealing the evolutionary history of the pSK1 family. Although improving, there remains a substantial need for bioinformatic software and pipelines that can accomplish such analyses in a computationally efficient manner. That said, sequence analysis alone would not have revealed the impact that pSK1 evolution has had on ST239 MRSA. That required detailed phenotypic profiles for antibiotic and disinfectant susceptibility; the latter necessitating the development of a modified method to discern the subtle difference in susceptibility between populations. Thus, this work also serves as a framework for how plasmid evolution can be extensively studied. One of the troubling findings of this work is that there has been a population level shift towards reduced susceptibility to biocides, suggesting that biocide use in healthcare environments may be factor influencing the evolution of *S. aureus*, and likely other pathogenic bacteria. This raises questions about the need for regulations around biocide use to prevent the emergence of highly tolerant or biocide resistant strains.

Chapter 5

Discussion

5.1 Introduction

Antibiotic resistant pathogens pose a significant threat to human health and *S. aureus* is a major cause of disease in humans. Central to the prevention and management of staphylococcal disease is a deep and comprehensive understanding of the bacterium. The body of work presented in this thesis augments and complements our understanding of the evolutionary pathways and molecular tools employed by highly successful lineages of antibiotic-resistant *S. aureus*, thus enabling them to evade measures of disease prevention, management, and eradication. Three examples have been presented, each detailing a different aspect of staphylococcal evolution: (i) persistence of the HA-MRSA lineage ST239 in Australian hospitals; (ii) rapid emergence of FA-resistance ST5 MRSA in New Zealand; and (iii) the evolution of the pSK1 plasmid family and the influence that has had antibiotic and biocide susceptibility in ST239 MRSA. The following chapter will illustrate how these three pieces of work fit together and the insight they provide in understanding the evolutionary dynamic of successful lineages of MRSA in Australia and New Zealand.

5.2 Key Findings & Significance

A central theme to the research presented in this thesis has been understanding ‘population dynamics’, in terms of what evolutionary factors contribute to the emergence and persistence of MRSA lineages; what adaptive pathways and molecular tool are utilised and under what circumstances; what environmental factors influence adaptation; and what are the consequences, being intended or unintended, of these population level changes. While the work presented in this thesis has focused on what is occurring in Australia and New Zealand, the findings presented mirror those reported from other regions and common trends are apparent when this work is considered alongside the rapidly growing compendium of population-based bacterial genomic studies.

5.2.1 Emergence & Establishment of Novel Lineages

From a Darwinian perspective, the ability for a given population to outcompete and establish itself in a given niche is driven by possession of an adaptive advantage over other populations. In terms of antibiotic resistant bacteria, we observe this as either the gain or development of a resistance mechanism(s) that provides a survival advantage in an antibiotic-exposed environment. While the benefit from such adaptation represents a spectrum, in this context we consider the emergence of a ‘novel lineage’ to represent an ancestral lineage that has undergone a significant adaptive shift or one that has been introduced into a new environment for which it already comprises a survival advantage; both of which have been observed in the staphylococcal populations investigated in this thesis.

Lineage emergence through gene acquisition

In the case of the AK3 (ST5) MRSA clone in New Zealand, we observed a significant adaptive shift in a local ancestral ST5 MSSA lineage; the gain of a novel chimeric *SCCmecIVa-fusC* II to create a FA-resistant MRSA lineage with a significant survival advantage in a region with high population level use of a topical antimicrobial. The success of AK3 has further been supplemented by the high local incidence of staphylococcal SSTI; the consequence of local cultural and socio-economic factors, and a population comprised of higher proportions of ethnicities that are known to be disproportionately affected by staphylococcal disease (22). However, the impact of FA use on influencing staphylococcal evolution is not limited to New Zealand or ST5 MRSA, with multiple novel FA-resistant staphylococcal lineages have been reported in geographically distinct regions (47, 193, 346, 347, 383). The increasing use of genomics in these studies paints a picture of heightened genomic plasticity in the associated MGEs, allowing for the development of multiple novel *fusC*-harbouring SCC and *SCCmec* elements which are being acquired by diverse staphylococcal lineages in response to the evolutionary pressure imposed by FA use (193, 194, 228). While we have described six *fusC*-harbouring SCC and *SCCmec* elements in Chapter 3, FIG 3 in (194) and Table 3.1, we have identified a further five novel structures carried by ST1, ST5 and ST8 *S. aureus* (unpublished data). Most *fusC*-harbouring staphylococcal clones have not been as outwardly successful as the ST5 and ST1 lineages in New Zealand, a likely consequence of the unique environment in which these clones have emerged. This work further highlights that changes in that environment, namely the implementation of regulations to FA use, could significantly impact the ongoing evolution of these lineages and the emergence of new FA-resistant clones by abrogating the adaptive advantage provided by *fusC*; the effect of which would likely mirror what was observed in New Zealand with unregulated mupirocin use and the subsequent decrease in the prevalence of mupirocin-resistant staphylococci following regulation of the topical antimicrobial (39), https://surv.esr.cri.nz/PDF_surveillance/Antimicrobial/Staph/Staph_1999.pdf).

In the case of ST239, we observed a similar situation early in its evolutionary history linked with the appearance of the pSK1 plasmids. We have estimated that the earliest acquisition event for an pSK1-like plasmid into the Australian ST239 clade occurred during the late 1970s and in short succession a change in plasmid configuration, specifically acquisition of an aminoglycoside-resistance conferring transposon, *Tn4001* [FIG 3 in (204)]; this occurrence coinciding with early clinical use of gentamicin in Australia. Although speculative, it follows that the acquisition of a gentamicin-resistance mechanism would provide an adaptive advantage for ST239 MRSA, the pSK1 plasmid serving as vehicle to assist in the mobilisation of the required genetic material and could explain the increase in pSK1 carriage observed during this time [FIG 1B in (204)]. Unfortunately, the limited availability of historic isolate collections when combined with the known limitation of the typing techniques utilised at the time (predominantly phage-typing) and their discrepancies with MLST and WGS, the reported incidence of gentamicin-resistance in ST239 and non-ST239 MRSA is imprecise.

Lineage emergence through introduction into a new environment

Later in the evolution of ST239 MRSA in Australia, we observed an example of lineage emergence through the introduction of an established lineage into a new environment. This event went undetected because the introduced clone, the Asian-Australian ST239 clade, represented the same sequence type as the predominant local lineage, highlighting an increasingly acknowledged flaw in traditional MLST; even in species such as *S. aureus*, where

vertical descent is the primary mode of evolution, MLST can be insufficiently discriminant in dividing populations for the purpose of surveillance. Further hindering the identification of the Asian-Australian ST239 clade, was the fact that the suspected adaptive advantage possessed by this population was subtle, a slight but significant shift in susceptibility to anti-MRSA antibiotics (the glycopeptides and daptomycin). This was further obscured by the Australian ST239 clade which was convergently evolving the same phenotype. Additionally, highlighting that use of these agents is strongly influencing ST239 MRSA evolution. While the two clades could be distinguishable by PFGE (unpublished data), a genomic investigation of geographically diverse dataset with substantial representation of Australian isolates was required to identify that the Asian-Australian clade originated in South-East Asia and its appearance in Australia was the result of intercontinental transmission and not adaptation of the local Australia clade. This knowledge is important in the context of designing and implementing programs aimed at reducing antibiotic-resistant bacterial infections by precluding transmission.

5.2.2 Adaptation to Promote Lineage Persistence

The success of a lineage to persist in a given environment is facilitated by its ability to continually adapt in a way that maintains its competitive advantage over other populations, a feature strongly supported by the underlying genetic diversity of the population, allowing for plasticity. In the case of staphylococcal lineages that persist in healthcare environments, their survival advantage has largely been attributed to their success at developing an enhanced and varied resistome (19). Like most of the traditional HA-MRSA lineages, since its emergence ST239 has been defined by its multiply antibiotic resistant phenotype and our investigations demonstrated that this population is continuing to augment this characteristic in response to extended exposure to healthcare environments (137). Multiple evolutionary mechanisms that promote persistence have been identified in this work and broadly can be classified into two groups, those that facilitate an increase in population genetic diversity and those which represent adaption to a specific niche through positive selection and genome refinement.

Gene acquisition to maintain and/or increase population genetic diversity

The introduction of the Asian-Australian ST239 clade in to Australia, in addition to establishing a new lineage in the region had the potential to significantly boost the genetic diversity of the collective ST239 gene pool. Further, it could have provided redundancy for the resistome, both in the encoded mechanisms of resistance and the MGE which facilitate their movement, as the two clades have distinct genotypes; establishing a situation in which the most advantage and non-deleterious combination of resistance-associated genes and MGEs could be combined. However, there was no evidence to support that these clades shared genetic material while in Australia, likely resulting from their state-level separation and negating some of the benefit of this increase in diversity. The modelling undertaken to assess changes in effective population size (FIG 2A in (137), and Figure 2.5) suggested that the introduction of the Asian-Australian clade has maintained the collective population's diversity but not substantially increased it, a finding consistent with the phenomenon of clonal replacement. That is not to say these clades have not shared genetic material or been influenced by non-ST239 lineages. A key limitation of the work presented in this thesis is that it has focused on single lineages in isolation and as a result would have missed any crucial evolutionary events that involved other lineages. For example, the donor of the SCCmecIVa-*fusC* II carried by the ST5 MRSA lineage, in addition to many of the other *fusC*-harbouring SCC elements identified in *S. aureus*,

remain unknown. Subsequently, it is unclear whether the chimeric or composite structures are evolving in the FA-resistant *S. aureus* lineages in which they have been identified or in an alternative host, potentially a member of the skin microbiota. If it is the latter, then the consequences of unregulated FA may be much greater than what is observed when only cases of staphylococcal infection are considered. Future studies which are likely to provide the greatest insight into staphylococcal population dynamics are going to be those which consider all *S. aureus* or all staphylococci in a given environment, whether that be through examining diverse collection of isolates using similar approaches to those described in this thesis or those techniques being developed in the field of metagenomics.

Increasing genetic diversity through co-selection

Molecular and genomic investigations of antibiotic-resistance bacteria have clearly demonstrated that resistance-associated genes frequently move together, co-located on a common MGE (54, 104). Therefore, through the phenomenon of co-selection, the adaptive advantage provided by one gene may promote the acquisition of an MGE and consequently all the genes that it carries, provided co-carriage does not abrogate that adaptive advantage. Thus, co-selection serves as another mechanism by which to introduce genetic diversity into a population and support genome plasticity. Both MGEs investigated in this thesis are examples of this. The prototypic pSK1 plasmid carries three genes encoding unrelated resistance mechanisms, *aac(6')-aph(2'')*, *drfA*, and *qacA*, with other reported pSK1 plasmid variants known to carry additional resistance-associated genes such as *blaZ* (204, 288). Co-located in SCCmecIVa-*fusC* II are *fusC*, *mecA* and *tirS*. The latter, a putative virulence factor hypothesised to be involved in immune evasion and adhesion (194, 250). The expression of *tirS* has been shown *in vitro* to be augmented in the presence of FA, and the gene is invariably carried in all identified *fusC*-harbouring SCC and SCCmec elements (193, 194). Unlike the structural changes observed with pSK1, the J1 region of the *fusC*-harbouring SCC and SCCmec elements (where both *fusC* and *tirS* are located) is highly conserved, as illustrated in FIG 3 in (194). It is unclear whether this region is extremely stable compared to the rest of the MGE or the two genes are beneficial to one another beyond their individual functions; a question that could not be answered by this work. In general, the greatest consequence of this co-location is that single acquisition event can create novel staphylococcal lineages that are multidrug resistance with enhanced pathogenetic potential. This is an area that requires further investigation, both in understanding the biology facilitating the co-location of these genes but also the impact of antibiotic use in the context of co-selection and how antibiotic stewardship can be tailored to mitigate this.

Positive selection of resistance-associated genotypes

As a population evolves it will shift to reflect the genotype of its most adapt members. This process is evident in the Australian ST239 clade in respect to the identified shifts in glycopeptide and daptomycin susceptibility, and chlorhexidine tolerance; all phenotypes that provide an adaptive advantage in the healthcare environment. The work presented in Chapter 2 alluded to the accumulation of mutations in the two-component global regulatory system *walKR* and the RNA polymerase, specifically the β -subunit, as likely contributing to the population level shift in glycopeptide susceptibility, FIG 3 in (137). This work is one of a limited number of studies to demonstrate the phenomenon of “MIC creep” in a single, genetically-defined clone. Since this study, there has been extensive bioinformatic development in the area of statistical genomics, prosperously applied to identified genotypes

associated with antimicrobial resistance. Revisiting this dataset utilising these new bioinformatics tools may enable further exploration of other mutations contributing to reduced glycopeptide and daptomycin susceptibility in the ST239 population.

In the case of chlorhexidine tolerance, the work presented in Chapter 4 demonstrated through multiple approaches the association between enhanced tolerance and structural changes in the configuration of the pSK1 plasmids, notably chromosomal integration and large deletions in the plasmid backbone invariably affecting the plasmid replication machinery. While the latter would be strongly influenced by the need to stabilise the plasmid sequence in the chromosome by removing incompatibility with the native replication machinery, the effect of both changes may have altered the regulation of *qacA*; a hypothesis that requires further exploration. What is clear is that there was a strong selection pressure favouring chromosomally integrated plasmid variants, evident from the multiple independent integration events identified in the Australian ST239 clade, illustrated in FIG 3 in (204).

Consequences of adaptation: trading virulence for resistance

In the investigation of the ST239 MRSA clades, it was identified that while these lineages were evolving enhanced antibiotic resistance and disinfectant tolerance, they were concurrently evolving reduced replicative fitness and virulence potential [TABLE 1 and FIG 4 in (137)]. It is unclear whether this is a consequence of the mechanisms of resistance developed or acquired in ST239 MRSA or if this phenotype also provides an adaptive advantage that has undergone positive selection. It is likely a combination of both. Mutational-based resistance to the glycopeptides, such as those involving *walkR* and *rpoB* are known to impact on fitness and have been demonstrated in both *in vitro* and *in vivo* models (145, 155, 305). However, there is also a growing body of research that suggests that *agr*-dysfunction, occurring through both altered regulation and disruption of the *agr* locus through mutation, may actually provide a survival advantage in certain host environments. It is increasingly being recognised that *agr*-dysfunction is a common phenotype amongst staphylococcal isolates recovered from healthcare environments (137, 274, 275, 323, 324, 384), leading to the hypothesis that naturally occurring *agr*-inactive mutants undergo positive selection in this environment; these isolates demonstrating reduced immunogenicity and an association with persistent infection (274). Further, *agr* activity has been exhibited to impose a fitness cost on the bacterium when exposed to certain antimicrobials (385). However, evolution of a population towards *agr*-dysfunction, while potentially advantageous during infection, is ultimately a limiting pathway as this phenotype has been shown to be disadvantageous outside this environment, with *agr*-inactive mutants readily removed from natural staphylococcal populations (275).

5.2.3 Lineage Disappearance

An aspect of bacterial population dynamics that is largely understudied and has not been considered in this thesis are the circumstances surrounding the disappearance of a lineage. Over the last two decades, the ST239 MRSA population in Australia has been in steady decline and as of 2018 represents only 16.2% of HA-MRSA isolates from national surveillance of SAB (<http://agargroup.org.au/agar-surveys#Staphylococcus-aureus>). Having said that, following the global ST22 HA-MRSA clone, also known as EMRSA-15, ST239 remains the second most commonly recovered HA-MRSA lineage in the region. The replacement of ST239 MRSA by EMRSA-15 has

been observed in several countries over the last few decades, with multiple studies reporting data in support of EMRSA-15 having a competitive advantage over ST239 (386-389). However, there is also data suggesting that the decline in ST239 MRSA began before the expansion of EMRSA-15 and although a clonal replacement event has occurred, the incidence of EMRSA-15 is substantially less than that of ST239 MRSA (303). This finding strongly suggests that there are other factors contributing to the decline of ST239 MRSA. These are likely to include changes in the healthcare environment over the last few decades, such as the implementation of improved infection control measures and hand hygiene initiatives, changes in protocols to reduce patient's length of stay, and the implementation and improved adherence to antimicrobial stewardship programs. Further contributing to the phenomenon of lineage disappearance is the concept that the evolutionary pathway ST239 MRSA has taken to adapt to the highly selective healthcare environment has ultimately restricted its ability to survive outside of it. Future research efforts should explore these dynamics to improve our comprehension about what factors have the greatest influence on lineage disappearance and how these can be exploited to reduce the burden of antibiotic-resistant staphylococcal infection.

Appendix 1

Chapter 2: Supplementary Datasets

Publication: Dataset S1

Extended Analysis: Supplementary Dataset

Dataset S1 - Summary of study isolates and genome sequencing

Study ID	Year	Country	Institution, State	Site	Comments	Reference	ST,Method	WGS Platform	Fold Coverage ^A
BPH0468	2005	Australia	QHP, Queensland	Tissue	Daptomycin NS	Kelley, JAC, 2011	ST239 - MLST	MiSeq	48.60
BPH0475	2003	Australia	RMH, Victoria	Blood Culture	hVISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	77.27
BPH0488	2003	Australia	RMH, Victoria	Blood Culture	hVISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	62.35
BPH0491	2005	Australia	STV, Victoria	Blood Culture	hVISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	52.35
BPH0494	2005	Australia	QHP, Queensland	Blood Culture	hVISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	43.98
BPH0496	2004	Australia	QHP, Queensland	Tissue	hVISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	42.07
BPH0499	2005	Australia	MEP, Victoria	Blood Culture	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	46.50
BPH0502	2003	Australia	New South Wales	Tissue	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	54.56
BPH0503	2003	Australia	Victoria	Blood Culture	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	52.96
BPH0508	2006	Australia	QHP, Queensland	Blood Culture	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	53.75
BPH0509	2005	Australia	QHP, Queensland	Sputum	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	37.73
BPH0510	2005	Australia	MAH, New South Wales	Blood Culture	VISA	Kelley, JAC, 2011	ST239 - MLST	MiSeq	70.71
BPH2002	2002	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	42.26
BPH2003	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	55.33
BPH2004	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	47.48
BPH2005	1999	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	63.90
BPH2006	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	47.61
BPH2007	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	57.89
BPH2008	2000	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	75.26
BPH2009	2000	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	62.32
BPH2010	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	41.55
BPH2011	2001	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	36.01
BPH2012	1999	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	61.00
BPH2013	2004	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	40.22
BPH2014	1999	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	49.19
BPH2015	2004	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	45.21
BPH2016	2005	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	66.54
BPH2017	2005	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	40.46
BPH2018	2004	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	69.07
BPH2019	2005	Australia	MEH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	68.19
BPH2021	2004	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	51.29
BPH2022	2005	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	47.83
BPH2023	2006	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	38.39
BPH2024	2007	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	50.02
BPH2025	2008	Australia	AUH, Victoria	Blood Culture		Holmes, JID, 2011	ST239 - PCR	MiSeq	47.44
BPH2026	2007	Australia	MMC, Victoria	Blood Culture		Holmes, JID, 2011	ST239 - PCR	MiSeq	40.85
BPH2028	2007	Australia	MMC, Victoria	Blood Culture		Holmes, JID, 2011	ST239 - PCR	MiSeq	57.35
BPH2029	2009	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	59.73
BPH2031	2008	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	47.72
BPH2032	1998	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	51.32
BPH2033	2008	Australia	AUH, Victoria	Blood Culture		Holmes, JID, 2011	ST239 - PCR	MiSeq	72.27

BPH2034	2008	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - MLST	MiSeq	68.16
BPH2035	2008	Australia	AUH, Victoria	Blood Culture		Holmes, JID, 2011	ST239 - PCR	MiSeq	77.30
BPH2037	2008	Australia	AUH, Victoria	Blood Culture	MSSA	Holmes, JID, 2011	ST239 - PCR	MiSeq	45.56
BPH2045	2007	Australia	WEH, New South Wales	Blood Culture	MSSA	Holmes, JID, 2011	ST239 - PCR	MiSeq	40.45
BPH2054	1980	Australia	ALH, Victoria			This Study	ST239 - PCR	MiSeq	58.51
BPH2055	1980	Australia	ALH, Victoria			This Study	ST239 - PCR	MiSeq	48.64
BPH2056	1981	Australia	ALH, Victoria			This Study	ST239 - PCR	MiSeq	71.83
BPH2057	1981	Australia	ALH, Victoria			This Study	ST239 - PCR	MiSeq	43.96
BPH2059	1981	Australia	ALH, Victoria		Freeze dried in 1981	This Study	ST239 - PCR	MiSeq	50.86
BPH2068	2009	Australia	AUH, Victoria			This Study	ST239 - MLST	MiSeq	49.90
BPH2070	1998	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	46.49
BPH2071	1998	Australia	AUH, Victoria	Blood Culture		This Study	ST239 - PCR	MiSeq	53.01
BPH2072	1982	Australia	RNS, New South Wales	Tissue		This Study *	ST239 - PCR	MiSeq	44.29
BPH2073	1982	Australia	RNS, New South Wales	Tissue		This Study *	ST239 - PCR	MiSeq	52.52
BPH2075	1982	Australia	RNS, New South Wales	Tissue		This Study *	ST239 - PCR	MiSeq	48.92
BPH2077	1982	Australia	RNS, New South Wales	Tissue		This Study *	ST239 - PCR	MiSeq	52.00
BPH2078	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	51.97
BPH2079	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	47.73
BPH2080	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	58.06
BPH2081	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	73.30
BPH2082	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	47.14
BPH2083	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	43.36
BPH2084	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	68.88
BPH2085	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	37.01
BPH2086	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	50.32
BPH2088	1982	Australia	RPA, New South Wales			This Study *	ST239 - PCR	MiSeq	57.01
BPH2090	2012	Australia	AUH, Victoria		MSSA	This Study	ST239 - PCR	MiSeq	62.75
BPH2094	1998	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	169.55
BPH2095	2003	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	84.28
BPH2096	2005	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	86.54
BPH2097	2007	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	76.54
BPH2098	2007	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	96.24
BPH2099	2007	Australia	LPH, New South Wales	Blood Culture		This Study **	ST239 - WGS	Ion Torrent PGM	206.32
JKD6000	2002	Australia	MMC, Victoria	Blood Culture	VSSA variant of JKD6001	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	129.45
JKD6001	2003	Australia	MMC, Victoria	Blood Culture	VISA variant of JKD6000	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	186.63
JKD6004	2002	Australia	RBH, Queensland	Blood Culture	hVISA variant of JKD6005	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	177.92
JKD6005	2002	Australia	RBH, Queensland	Blood Culture	VISA variant of JKD6004	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	200.37
JKD6007	2002	Australia	RBH, Queensland	Blood Culture	Morphological variant of JKD6004	This Study	ST239 - PCR	MiSeq	44.63
JKD6008	2003	New Zealand		Blood Culture	VSSA variant of JKD6009	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	51.08
JKD6009	2003	New Zealand		Tissue	VISA variant of JKD6008	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	236.32
JKD6021	2001	Australia	AUH, Victoria	Blood Culture	VSSA variant of JKD6023	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	95.70
JKD6022	2001	Australia	AUH, Victoria	Blood Culture	Morphological variant of JKD6023	This Study	ST239 - PCR	MiSeq	37.57
JKD6023	2001	Australia	AUH, Victoria	Blood Culture	VISA variant of JKD6021	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	144.36
JKD6051	2003	Australia	RBH, Queensland	Blood Culture	VSSA variant of JKD6052	Howden, AAC, 2006	ST239 - WGS	Genome Analyser Ilx	134.08

JKD6052	2004	Australia	RBH, Queensland	Blood Culture	hVISA variant of JKD6051	Howden, AAC, 2006	ST239 - WGS	Genome Analyser IIx	146.30
JKD6121	2005	Australia	AUH, Victoria	Blood Culture	SCV	This Study	ST239 - WGS	Genome Analyser IIx	181.09
BPH0365	1996	Turkey				Harris, Science, 2010	ST239 - PCR	MISeq	56.49
BPH0366	1996	Hungary				Harris, Science, 2010	ST239 - PCR	MISeq	65.53
2A8	2001	Czech Republic			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	34.58
AGT1	1997	Argentina			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	30.10
AGT9	1997	Argentina			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	29.79
AGT67	1997	Argentina			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	34.03
AGT120	1998	Argentina			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	29.20
BRA2	1997	Brazil			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	31.51
BZ48	1997	Brazil			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	28.05
CHI59	1998	China			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	27.00
CHI61	1998	China			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	26.54
CHL151	1998	Chile			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	28.38
DEN907	2001	Denmark			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	31.86
HGSA142	2003	Portugal			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	32.52
HSJ216	1997	Portugal			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	25.32
ICP5011	1993	Portugal			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	38.78
ICP5014	1993	Portugal			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	41.82
LHH1	1994	United States			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	35.78
S2	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	29.05
S7	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	54.89
S25	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	38.74
S26	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	46.72
S38	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	29.05
S39	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	31.59
S40	2006	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	31.59
S71	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	37.01
S78	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	31.34
S81	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	30.54
S87	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	27.25
S93	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	26.57
S102	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	26.24
S106	2007	Thailand			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	34.13
TUR1	1996	Turkey			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	25.57
TW20	2003	United Kingdom			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	-
URU34	1997	Uruguay			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	33.47
URU110	1998	Uruguay			WGS Reads Only	Harris, Science, 2010	ST239 - WGS	Genome Analyser IIx	27.33

* Strains provided courtesy of author NF; **Strains and WGS provided courtesy of authors SOJ and SJH; ^AMapped to TW20; Manufacturers of sequencing platforms used: Miseq (Illumina), Ion Torrent PGM (Life Technologies), Genome Analyser IIx (Illumina); Abbreviations: MSSA = methicillin-resistant *S. aureus*, NS = non susceptible, SCV = small colony variant, VISA = vancomycin-intermediate *S. aureus*, hVISA = heterogeneous VISA, WGS = Whole Genome Sequencing

Supplementary Dataset

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
1	BMB9393	-	1993	Brazil	-	Complete Genomes	454 GS FLX Titanium	CP005288, CP005289	-	Costa (2013) PMID:23929475	Costa. 2013. GenomeA. PMID:23929475
2	TW20	-	2003	United Kingdom	-	Complete Genomes	-	FN433596, FN433597, FN433598	-	Edgeworth (2007) PMID:17243050	Holden (2010) PMID: 19948800
3	Z172	-	2010	Taiwan	-	Complete Genomes	Illumina/Solexa + PacBio	CP006838, CP006839, CP006840	-	Chen (2013) PMID:24309740	Chen (2013) PMID:24309740
4	T0131	-	2006	China	-	Complete Genomes	Roche 454 / Solexa	CP002643	-	Li (2011) PMID:21551295	Li (2011) PMID:21551295
5	JKD6008	-	2006	New Zealand	-	Complete Genome	-	CP002120	-	Howden (2006) PMID:16940100	Howden (2010) PMID:20802046
6	XN108	-	-	China	-	Complete Genomes	Ion Torrent PGM	CP007447	-	Zhang (2014) PMID:25059856	Zhang (2014) PMID:25059856
7	GV69	-	1996	Brazil	-	Complete Genomes	454 GS FLX Titanium	CP009681; SAMN03144721	SRR2601051	Botelho (2016) PMID:27152133	Botelho (2016) PMID:27152133
8	HC1335	-	2001	Brazil	-	Complete Genomes	454 GS FLX Titanium	CP012012	-	Botelho (2016) PMID:27635055	Botelho (2016) PMID:27635055
9	M92	-	-	Canada	-	Complete Genomes	PacBio RS-II	CP015447	-	McClure (2017) PMID:28596402	McClure (2017) PMID:28596402
10	BPH0468	-	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212938	ERR732815	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
11	BPH0475	-	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212931	ERR732816	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
12	BPH0488	-	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212956	ERR732817	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
13	BPH0491	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212955	ERR732818	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
14	BPH0494	-	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212940	ERR732819	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
15	BPH0496	-	2004	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212928	ERR732820	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
16	BPH0499	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212937	ERR732821	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
17	BPH0502	-	2003	Australia	WA	WGS Reads	Illumina MiSeq	SAMEA3212919	ERR732822	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
18	BPH0503	-	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212967	ERR732823	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
19	BPH0508	-	2006	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212893	ERR732824	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
20	BPH0509	-	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212911	ERR732825	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
21	BPH0510	-	2005	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212957	ERR732826	Kelley. 2011. JAC. PMID:21393156	Baines (2015) PMID: 25736880
22	BPH2002	-	2002	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212904	ERR732827	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
23	BPH2003	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212913	ERR732828	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
24	BPH2004	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212949	ERR732829	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
25	BPH2005	-	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212961	ERR732830	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
26	BPH2006	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212925	ERR732831	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
27	BPH2007	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212952	ERR732832	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
28	BPH2008	-	2000	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212918	ERR732833	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
29	BPH2009	-	2000	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212948	ERR732834	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
30	BPH2010	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212899	ERR732835	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
31	BPH2011	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212962	ERR732836	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
32	BPH2012	-	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212958	ERR732837	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
33	BPH2013	-	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212934	ERR732838	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
34	BPH2014	-	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212969	ERR732839	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
35	BPH2015	-	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212954	ERR732840	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
36	BPH2016	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212908	ERR732841	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
37	BPH2017	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212936	ERR732842	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
38	BPH2018	-	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212942	ERR732843	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
39	BPH2019	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212951	ERR732844	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
40	BPH2021	-	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212903	ERR732845	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
41	BPH2022	-	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212945	ERR732846	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
42	BPH2023	-	2006	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212915	ERR732847	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
43	BPH2024	-	2007	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212944	ERR732848	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
44	BPH2029	-	2009	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212905	ERR732852	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
45	BPH2031	-	2008	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212965	ERR732853	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
46	BPH2032	-	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212926	ERR732854	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
47	BPH2034	-	2008	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212917	ERR732856	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
48	BPH2054	-	1980	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212883	ERR732860	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
49	BPH2055	-	1980	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212885	ERR732861	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
50	BPH2056	-	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212902	ERR732862	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
51	BPH2057	-	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212889	ERR732863	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
52	BPH2059	-	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212897	ERR732864	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
53	BPH2068	-	2009	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212906	ERR732865	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
54	BPH2070	-	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212890	ERR732866	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
55	BPH2071	-	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212888	ERR732867	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
56	BPH2072	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212947	ERR732868	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
57	BPH2073	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212927	ERR732869	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
58	BPH2075	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212896	ERR732870	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
59	BPH2077	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212932	ERR732871	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
60	BPH2078	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212950	ERR732872	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
61	BPH2079	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212922	ERR732873	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
62	BPH2080	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212929	ERR732874	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
63	BPH2081	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212941	ERR732875	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
64	BPH2082	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212920	ERR732876	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
65	BPH2083	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212886	ERR732877	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
66	BPH2084	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212892	ERR732878	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
67	BPH2085	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212907	ERR732879	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
68	BPH2086	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212900	ERR732880	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
69	BPH2088	-	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212968	ERR732881	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
70	BPH2090	-	2012	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212943	ERR732882	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
71	BPH2094	-	1998	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212946	ERR732883	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
72	BPH2095	-	2003	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212887	ERR732884	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
73	BPH2096	-	2005	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212923	ERR732885	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
74	BPH2097	-	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212939	ERR732886	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
75	BPH2098	-	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212884	ERR732887	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
76	BPH2099	-	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212933	ERR732888	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
77	JKD6000	-	2002	Australia	VIC	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212909	ERR732889	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
78	JKD6001	-	2003	Australia	VIC	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212898	ERR732890	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
79	JKD6004	-	2002	Australia	QLD	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212901	ERR732891	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
80	JKD6005	-	2002	Australia	QLD	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212895	ERR732892	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
81	JKD6007	-	2002	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212912	ERR732893	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
82	JKD6009	-	2003	New Zealand	NZ	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212894	ERR732894	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
83	JKD6021	-	2001	Australia	VIC	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212953	ERR732895	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
84	JKD6022	-	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212914	ERR732896	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
85	JKD6023	-	2001	Australia	VIC	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212964	ERR732897	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
86	JKD6051	-	2003	Australia	QLD	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212960	ERR732898	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
87	JKD6052	-	2004	Australia	QLD	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212916	ERR732899	Howden (2006) PMID:16940100	Baines (2015) PMID: 25736880
88	JKD6121	-	2005	Australia	VIC	WGS Reads	Illumina Genome Analyzer Iix	SAMEA3212963	ERR732900	Baines (2015) PMID: 25736880	Baines (2015) PMID: 25736880
89	2A8	-	2001	Czech Republic	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029549	ERR064912	Melter (2003) PMID: 14605130	Castillo-Ramirez (2012) PMID:23270620
90	3HK	-	2000	Czech Republic	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029566	ERR064907	Melter (2003) PMID: 14605130	Castillo-Ramirez (2012) PMID:23270620
91	AGT1	-	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029527	ERR064926	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
92	AGT9	-	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029509	ERR064923	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
93	AGT67	-	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029555	ERR064924	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
94	AGT120	-	1998	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029554	ERR064919	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
95	ANS46	-	1982	Australia	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029537	ERR064898	Dubin (1992) PMID: 1313490	Castillo-Ramirez (2012) PMID:23270620
96	BK2421	-	1996	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029510	ERR064899	Kreiswirth (1993) PMID: 8093647	Castillo-Ramirez (2012) PMID:23270620
97	BRA2	-	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029514	ERR064914	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
98	BRA36	-	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029559	ERR064917	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
99	BZ48	-	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029508	ERR064918	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
100	CHI59	-	1998	China	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029524	ERR064937	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
101	CHI61	-	1998	China	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029538	ERR064938	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
102	CHL1	-	1997	Chile	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029526	ERR064915	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
103	CHL151	-	1998	Chile	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029539	ERR064916	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
104	D71	-	1996	Germany	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862562	ERR024980	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
105	D90	-	1996	Germany	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862585	ERR025801	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
106	DEN907	-	2001	Denmark	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029548	ERR064948	Faria (2005) PMID: 15815005	Castillo-Ramirez (2012) PMID:23270620
107	DEU2	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862577	ERR026652	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
108	DEU3	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862574	ERR026653	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
109	DEU5	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862580	ERR026655	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
110	DEU6	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862581	ERR026656	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
111	DEU8	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862579	ERR026658	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
112	DEU10	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862583	ERR026649	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
113	DEU11	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862554	ERR026650	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
114	DEU12	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862553	ERR026651	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
115	DEU14	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862555	ERR026664	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
116	DEU15	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862558	ERR026665	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
117	DEU16	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862557	ERR026666	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
118	DEU17	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862559	ERR026667	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
119	DEU19	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862561	ERR026668	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
120	DEU20	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862560	ERR026669	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
121	DEU23	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862541	ERR026662	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
122	DEU25	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862542	ERR026663	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
123	DEU27	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862536	ERR026677	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
124	DEU28	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862537	ERR026678	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
125	DEU29	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862538	ERR026679	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
126	DEU30	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862531	ERR026680	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
127	DEU35	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862662	ERR026683	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
128	DEU36	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862661	ERR026684	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
129	DEU37	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862660	ERR026674	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
130	DEU38	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862667	ERR026675	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
131	DEU39	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862666	ERR026676	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
132	DEU41	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862664	ERR026690	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
133	DEU42	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862655	ERR026691	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
134	DEU43	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862654	ERR026692	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
135	DEU46	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862570	ERR026694	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
136	DEU47	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862573	ERR026695	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
137	DEU49	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862567	ERR026697	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
138	DEU50	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862566	ERR026687	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
139	ES26	-	1996	Spain	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862530	ERR024981	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
140	FFP103	-	1990	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029533	ERR064932	Aires de Sousa (1998) PMID: 9705398	Castillo-Ramirez (2012) PMID:23270620
141	GRE4	-	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029513	ERR064904	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
142	GRE18	-	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029546	ERR064902	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
143	GRE108	-	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029521	ERR064936	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
144	GRE317	-	1999	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029564	ERR064903	Aires de Sousa (2003) PMID: 12517842	Castillo-Ramirez (2012) PMID:23270620
145	H202	-	2006	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862584	ERR025802	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
146	H211	-	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862528	ERR024972	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
147	H216	-	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862587	ERR025803	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
148	H24	-	2005	Egypt	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862527	ERR024973	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
149	H482	-	2007	Romania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862595	ERR024983	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
150	HDG2	-	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029561	ERR064930	Sanches (1996) PMID: 9158792	Castillo-Ramirez (2012) PMID:23270620
151	HGSA9	-	1997	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029534	ERR064921	Amorim (2002) PMID: 12523627	Castillo-Ramirez (2012) PMID:23270620
152	HGSA142	-	2003	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029517	ERR064922	Amorim (2007) PMID: 17626175	Castillo-Ramirez (2012) PMID:23270620
153	HSA10	-	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029520	ERR064931	Amorim (2002) PMID: 12523627	Castillo-Ramirez (2012) PMID:23270620
154	HSA11	-	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029542	ERR064929	Amorim (2002) PMID: 12523627	Castillo-Ramirez (2012) PMID:23270620
155	HSJ216	-	1997	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029550	ERR064920	Aires de Sousa (1998) PMID: 9705398	Castillo-Ramirez (2012) PMID:23270620
156	HU106	-	1996	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029518	ERR064927	Oliveira (2001) PMID: 11158109	Castillo-Ramirez (2012) PMID:23270620
157	HU109	-	1996	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029567	ERR064909	Oliveira (2001) PMID: 11158109	Castillo-Ramirez (2012) PMID:23270620
158	HU4	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862618	ERR026705	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
159	HU5	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862625	ERR026706	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
160	HU6	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862624	ERR026707	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
161	HU7	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862623	ERR026708	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
162	HU8	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862622	ERR026709	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
163	HU9	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862627	ERR026710	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
164	HU11	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862634	ERR026701	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
165	HU13	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862636	ERR024668	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
166	HU14	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862637	ERR024672	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
167	HU15	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862630	ERR024673	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
168	HU16	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862631	ERR024674	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
169	HU17	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862632	ERR024675	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
170	HU21	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862639	ERR024678	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
171	HU23	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862643	ERR024679	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
172	HU25	-	1993	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029569	ERR024670	Teixeira (1995) PMID: 7494036	Castillo-Ramirez (2012) PMID:23270620
173	HU26	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862644	ERR024671	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
174	HUR18	-	1997	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029551	ERR064910	Oliveira (2001) PMID: 11158109	Castillo-Ramirez (2012) PMID:23270620
175	HUSA304	-	1993	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029531	ERR064928	de Lencastre (1997) PMID: 11864123	Castillo-Ramirez (2012) PMID:23270620
176	ICP5011	-	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029545	ERR064933	Couto (1995) PMID: 7559943	Castillo-Ramirez (2012) PMID:23270620
177	ICP5014	-	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029543	ERR064934	Couto (1995) PMID: 7559943	Castillo-Ramirez (2012) PMID:23270620
178	ICP5062	-	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029562	ERR064935	Couto (1995) PMID: 7559943	Castillo-Ramirez (2012) PMID:23270620
179	IU1	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862647	ERR024681	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
180	IU2	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862646	ERR024685	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
181	IU4	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862648	ERR024687	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
182	IU6	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862651	ERR024688	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
183	IU7	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862650	ERR024689	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
184	IU9	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862617	ERR024691	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
185	IU10	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862652	ERR024692	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
186	IU11	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862653	ERR024682	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
187	IU12	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862658	ERR024683	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
188	IU13	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862659	ERR024684	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
189	IU15	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862657	ERR024974	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
190	LHH1	-	1994	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029523	ERR064900	de Lencastre (1996) PMID: 9158794	Castillo-Ramirez (2012) PMID:23270620
191	LIT68	-	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862586	ERR025804	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
192	LIT76	-	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862593	ERR025805	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
193	LIT89	-	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862592	ERR025796	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
194	M116	-	2004	Vietnam	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862545	ERR024988	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
195	M278	-	2005	Portugal	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862548	ERR024991	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
196	M418	-	2006	India	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862549	ERR024992	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
197	M592	-	2006	Syria	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862550	ERR024993	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
198	M705	-	2007	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862613	ERR024958	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
199	M996	-	2008	China	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862551	ERR024994	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
200	M1229	-	2009	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862546	ERR024989	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
201	MAL1	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862552	ERR024984	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
202	MAL3	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862609	ERR024964	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
203	MAL9	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862607	ERR024966	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
204	MAL11	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862543	ERR024985	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
205	MAL119	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862611	ERR024962	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
206	MAL215	-	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862608	ERR024963	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
207	MU1	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862569	ERR026688	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
208	MU3	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862565	ERR026003	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
209	MU4	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862564	ERR026007	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
210	MU5	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862604	ERR026008	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
211	MU6	-	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862605	ERR026009	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
212	MU7	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862602	ERR026010	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
213	MU9	-	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862603	ERR026011	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
214	MU10	-	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862600	ERR026012	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
215	MU20	-	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862620	ERR026703	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
216	Na21	-	2006	Sri Lanka	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862544	ERR024986	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
217	NA32	-	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862615	ERR024967	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
218	P32	-	1996	Poland	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862589	ERR025795	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
219	R35	-	1987	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029532	ERR064901	Tilston (1989) PMID: 2543284	Castillo-Ramirez (2012) PMID:23270620
220	RA3	-	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862588	ERR025798	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
221	RA6	-	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862591	ERR025799	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
222	RA7	-	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862616	ERR024968	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
223	S2	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029563	ERR064953	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
224	S7	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029556	ERR064946	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
225	S21	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029536	ERR064958	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
226	S24	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029553	ERR064957	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
227	S25	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029535	ERR064951	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
228	S26	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029511	ERR064949	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
229	S38	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029568	ERR064947	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
230	S39	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029558	ERR064959	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
231	S40	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029528	ERR064944	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
232	S42	-	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029515	ERR064955	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
233	S71	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029541	ERR064943	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
234	S78	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029525	ERR064954	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
235	S81	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029565	ERR064956	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
236	S85	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029529	ERR064939	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
237	S87	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029512	ERR064941	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
238	S93	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029560	ERR064942	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
239	S97	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029522	ERR064950	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
240	S102	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029519	ERR064945	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
241	S106	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029547	ERR064952	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
242	S130	-	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029516	ERR064940	Nickerson (2009) PMID: 19652705	Castillo-Ramirez (2012) PMID:23270620
243	TUR1	-	1996	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029552	ERR064906	Aires de Sousa (2003) PMID: 12904393	Castillo-Ramirez (2012) PMID:23270620
244	TUR9	-	1995	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029530	ERR064905	Aires de Sousa (2003) PMID: 12734244	Castillo-Ramirez (2012) PMID:23270620
245	TUR27	-	1996	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029557	ERR064908	Aires de Sousa (2003) PMID: 12904393	Castillo-Ramirez (2012) PMID:23270620
246	UK102	-	1996	United Kingdom	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862590	ERR025800	Castillo-Ramirez (2012) PMID:23270620	Castillo-Ramirez (2012) PMID:23270620
247	URU34	-	1997	Uruguay	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029544	ERR064925	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
248	URU110	-	1998	Uruguay	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029540	ERR064911	Aires de Sousa (2001) PMID: 11376057	Castillo-Ramirez (2012) PMID:23270620
249	SR2852	-	2005	United States	-	WGS Reads	Illumina HiSeq 2000	SAMN02649763	RRR1177896	Richter (2011) PMID: 21976769	Alam (2014) PMID:24787619
250	031_01	DB025818/01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800273	ERR029457	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
251	362_05	DB046586/05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800274	ERR029458	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
252	058_08	DR007259/08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800253	ERR029464	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
253	352_05	DB044224/05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800280	ERR029467	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
254	131_09	DB070928/09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800276	ERR029469	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
255	009_10	DM017526/10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800287	ERR029470	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
256	134_09	DB071635/09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800277	ERR029472	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
257	135_09	DB072476/09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800278	ERR029473	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
258	049_05	DM007329/05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800292	ERR029476	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
259	CGH1_08	8011775603/08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800297	ERR030155	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
260	CGH1_10	0043716978/10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800177	ERR030157	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
261	CGH2_08	8031866998/08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800181	ERR030159	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
262	CGH3_10	0060306746/10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800213	ERR030170	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
263	CGH6_06	6102297266/06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800382	ERR030173	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
264	CGH4_09	9073350043/09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800215	ERR030176	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
265	CGH4_10	0080344437/10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800216	ERR030177	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
266	CGH5_09	9093428285/09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800383	ERR030182	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
267	012_00	DB057108/00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800207	ERR030186	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
268	011_06	DM020975/00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800205	ERR030191	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
269	003_06	DB001858/06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800206	ERR030192	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
270	106_06	DX0254/06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800200	ERR030194	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
271	100_06	DB10089/06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800201	ERR030195	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
272	106_07	DR005123/07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800360	ERR030200	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
273	339_07	DR015686/07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800357	ERR030201	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
274	157_04	DB037404/04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800333	ERR030202	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
275	019_01	DM02138/01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800332	ERR030203	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
276	002_00	DB040859/00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800331	ERR030204	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
277	020_01	DB006232/01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800329	ERR030205	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
278	101_07	DR004988/07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800327	ERR030209	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
279	030_05	DB019675/05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800258	ERR030213	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
280	255_04	AH255/04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800353	ERR030216	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
281	259_04	AH259/04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800354	ERR030217	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
282	263_04	AH263/04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800352	ERR030219	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
283	264_04	AH264/04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800261	ERR030220	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
284	028_01	DM07777/01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800260	ERR030221	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
285	069_02	DB065511/02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800271	ERR030224	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
286	006_00	DB043050/00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800285	ERR030228	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
287	067_02	DM020544/02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800265	ERR030232	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
288	068_02	DB065965/02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800272	ERR030233	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
289	071_03	DM001208/03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800269	ERR030235	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
290	078_05	DB022127/05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800264	ERR030238	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
291	086_03	DB057872/03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800263	ERR030239	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
292	009_00	DB040363/00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800336	ERR030240	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
293	TTSH1_01	1351-118	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800251	ERR030242	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
294	TTSH2_07	7033-089773	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800229	ERR030249	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
295	TTSH1_10	0063-543813	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800219	ERR030252	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
296	TTSH2_03	3038-49	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800226	ERR030255	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
297	TTSH2_08	8023-032840	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800230	ERR030259	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
298	TTSH3_07	7033-088525	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800196	ERR030260	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
299	TTSH3_08	8023-032825	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800232	ERR030261	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
300	TTSH3_03	3039-17	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800234	ERR030266	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
301	TTSH4_02	2006-106	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800293	ERR030276	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
302	TTSH4_03	3046-47	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800372	ERR030277	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
303	TTSH5_02	2008-12	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800345	ERR031666	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
304	TTSH5_06	6013-651744	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800346	ERR031670	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
305	WAU_8	WBG4997	1982	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021823	ERR053014	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
306	WAU_9	WBG6119	1985	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021807	ERR053015	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
307	WAU_10	WBG8484	1985	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021856	ERR053016	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
308	WAU_16	WBG8490	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1022012	ERR053022	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
309	WAU_17	WBG8491	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021973	ERR053023	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
310	WAU_18	WBG8492	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021952	ERR053024	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
311	WAU_19	WBG8493	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021934	ERR053025	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
312	WAU_21	WBG8495	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021910	ERR053027	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
313	WAU_26	WBG9016	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021849	ERR053032	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
314	WAU_29	WBG9019	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021995	ERR053035	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
315	WAU_31	WBG9022	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021970	ERR053037	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
316	WAU_34	WBG9023.3	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021933	ERR053040	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
317	WAU_39	WBG9028	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021951	ERR053044	Hsu (2015) PMID:25903077	Hsu (2015) PMID:25903077
318	T9_N3	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958822	ERR023888	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
319	T234_W2	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958806	ERR023872	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
320	T71_N2	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958988	ERR033836	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
321	T99_N5	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958963	ERR033828	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
322	T12_N1	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958848	ERR023884	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
323	T271_C1	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958810	ERR023867	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
324	T35_N2	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958775	ERR033823	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
325	T95_N2	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958962	ERR033833	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
326	T192_T3	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958900	ERR023907	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
327	T303_N3	-	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958836	ERR033080	Tong (2015) PMID:25491771	Tong (2015) PMID:25491771
328	SA13002	-	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03856026	SRR2097047	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
329	SA13005	-	2013	China	-	WGS Reads	Illumina MiSeq	SAMN03878399	SRR2106469	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
330	SA13007	-	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878400	SRR2106470	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
331	SA13009	-	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878401	SRR2106471	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
332	SA13012	-	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878402	SRR2106472	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
333	SA13023	-	2013	China	-	WGS Reads	Illumina MiSeq	SAMN03878403	SRR2106473	Kong (2016) PMID:26950298	Kong (2016) PMID:26950298
334	BSAC_st3021	-	2010	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464569	ERR124493	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
335	BSAC_s3046	-	2010	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464296	ERR124495	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
336	BSAC_st697	-	2003	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464583	ERR129279	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
337	D8	610	2007	Gambia	-	WGS Reads	Illumina MiSeq	SAMEA3727072	ERR1213802	Senghore (2016) PMID:27474712	Senghore (2016) PMID:27474712
338	DS_009	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448838	ERR1069915	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
339	DS_014	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448843	ERR1069920	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
340	US_002	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448943	ERR1070030	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
341	US_008	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448948	ERR1070035	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
342	US_030	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448965	ERR1070052	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
343	US_017	-	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448955	ERR1070042	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
344	BPH2065	-	1981	Australia	SA	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
345	BPH2066	-	1981	Australia	SA	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
346	BPH2104	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
347	BPH2106	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
348	BPH2107	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
349	BPH2108	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
350	BPH2109	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
351	BPH2113	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
352	BPH2119	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
353	BPH2120	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
354	BPH2139	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
355	BPH2154	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
356	BPH2157	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
357	BPH2161	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
358	BPH2163	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
359	BPH2165	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
360	BPH2167	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
361	BPH2168	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
362	BPH2196	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
363	BPH2198	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
364	BPH2199	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
365	BPH2202	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
366	BPH2203	-	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
367	BPH2204	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
368	BPH2205	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
369	BPH2207	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
370	BPH2210	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
371	BPH2213	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
372	BPH2219	-	1991	Australia	TAS	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
373	BPH2222	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
374	BPH2226	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
375	BPH2228	-	1990	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
376	BPH2230	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
377	BPH2232	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
378	BPH2233	-	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
379	BPH2234	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
380	BPH2235	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
381	BPH2236	-	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
382	BPH2238	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Baines (2019) PMID: 30783008	Baines (2019) PMID: 30783008
383	BPH2240	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
384	BPH2242	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
385	BPH2243	-	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Unpublished	Unpublished
386	BPH2710	-	2011	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
387	BPH2718	-	2011	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
388	BPH2723	-	2011	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
389	BPH2726	-	2011	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
390	BPH2747	-	2011	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
391	BPH2779	-	2011	Australia	VIC	WGS Reads	Illumina HiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
392	BPH2780	-	2011	Australia	VIC	WGS Reads	Illumina HiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
393	BPH2781	-	2011	Australia	VIC	WGS Reads	Illumina HiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
394	BPH2787	-	2011	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
395	BPH2798	-	2011	Australia	VIC	WGS Reads	Illumina HiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
396	BPH2824	-	2011	Australia	VIC	WGS Reads	Illumina HiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
397	BPH2827	-	2011	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
398	BPH2869	-	2012	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
399	BPH2896	-	2012	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
400	BPH2946	-	2012	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
401	BPH2947	-	2012	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
402	BPH3202	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
403	BPH3208	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
404	BPH3220	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
405	BPH3234	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
406	BPH3244	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
407	BPH3250	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
408	BPH3254	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
409	BPH3260	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
410	BPH3262	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
411	BPH3266	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
412	BPH3272	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
413	BPH3276	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
414	BPH3284	-	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
415	BPH3306	-	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
416	BPH3324	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
417	BPH3328	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
418	BPH3332	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
419	BPH3344	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
420	BPH3346	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
421	BPH3348	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
422	BPH3356	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
423	BPH3360	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
424	BPH3362	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
425	BPH3366	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
426	BPH3368	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
427	BPH3370	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
428	BPH3372	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
429	BPH3376	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
430	BPH3382	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
431	BPH3394	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
432	BPH3401	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
433	BPH3403	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
434	BPH3413	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
435	BPH3415	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
436	BPH3417	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
437	BPH3431	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
438	BPH3433	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
439	BPH3435	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
440	BPH3439	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
441	BPH3443	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
442	BPH3451	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
443	BPH3453	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
444	BPH3455	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
445	BPH3457	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
446	BPH3463	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
447	BPH3465	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
448	BPH3469	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
449	BPH3483	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
450	BPH3485	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
451	BPH3493	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
452	BPH3495	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
453	BPH3499	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
454	BPH3507	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
455	BPH3511	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
456	BPH3523	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
457	BPH3525	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
458	BPH3527	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
459	BPH3529	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
460	BPH3531	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
461	BPH3533	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
462	BPH3535	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
463	BPH3541	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
464	BPH3543	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
465	BPH3545	-	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
466	BPH3547	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
467	BPH3549	-	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
468	BPH3632	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished

	ID	Alternate ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
469	BPH3634	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
470	BPH3636	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
471	BPH3638	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
472	BPH3640	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
473	BPH3643	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
474	BPH3644	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
475	BPH3646	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
476	BPH3648	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
477	BPH3650	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
478	BPH3652	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
479	BPH3658	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
480	BPH3660	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
481	BPH3662	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
482	BPH3664	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
483	BPH3666	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
484	BPH3668	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
485	BPH3670	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
486	BPH3672	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
487	BPH3674	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
488	BPH3676	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
489	BPH3678	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
490	BPH3680	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
491	BPH3682	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
492	BPH3684	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
493	BPH3686	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
494	BPH3688	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
495	BPH3692	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
496	BPH3694	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
497	BPH3696	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
498	BPH3698	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
499	BPH3700	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
500	BPH3702	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
501	BPH3704	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
502	BPH3706	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
503	BPH3708	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
504	BPH3710	-	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
505	BPH3712	-	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
506	S20209-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
507	S20305-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
508	S20603-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
509	S20607-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
510	S20608-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
511	S20610-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
512	S20613-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
513	S20617-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
514	S20622-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
515	S20623-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
516	S20629-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
517	S20826-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
518	S23616-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
519	S23619-2015	-	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
520	S31902-2015	-	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
521	S31904-2015	-	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
522	S31913-2015	-	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
523	S31918-2015	-	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
524	S32313-2015	-	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
525	S41210-2015	-	2015	Australia	QLD	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
526	S51001-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
527	S51003-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
528	S51006-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
529	S51007-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
530	S51008-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
531	S51014-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
532	S51024-2015	-	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
533	S51408-2015	-	2015	Australia	SA	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
534	S51504-2015	-	2015	Australia	SA	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished

Appendix 2

Chapter 3: Supplementary Material

Publication: Supplementary Dataset S1

Extended Analysis: Supplementary Dataset

Supplementary Table 1. Demographics and genome sequencing information for all bacterial isolates used in this study

Isolate	Location	Year	MRSA/ MSSA	MLST*	Clonal Complex	Fusidic Acid Susceptibility	Sequencing Platform	WGS Format	Source [Reference]
502A	United States	1963	MSSA	5	5	-	-	Complete Assembly	SRA: CP007454-5 [9]
ECT-R2	Sweden	2004-7	MSSA	5	5	-	-	Complete Assembly	SRA: FR714927-9 [10]
ED98	United Kingdom	1996-7	MSSA	5	5	-	-	Complete Assembly	SRA: CP001791.1, CP001782-4 [11]
HOU1444-VR	Brazil	2012	MSSA	5	5	-	-	Complete Assembly	SRA: CP012593-6 [12]
JH1	United States	2000	MRSA	5	5	-	-	Complete Assembly	SRA: CP000736-7 [13]
Mic3	United States	1996	MRSA	5	5	-	-	Complete Assembly	SRA: CP000703-4 [13]
Mu50	Japan	1997	MRSA	5	5	-	-	Complete Assembly	SRA: AP009324.1 [14]
N315	Japan	1982	MRSA	5	5	-	-	Complete Assembly	SRA: BA000017.4, AP003367.1 [14]
SA564	-	-	MSSA	5	5	-	-	Complete Assembly	SRA: BA000018.3, AP003139.1 [8]
FCFHV36	Brazil	-	MRSA	105	5	-	-	Complete Assembly	SRA: CP010890-1 [15]
MSSA476 ^{a, b}	United Kingdom	1997	MSSA	1	1	-	-	Complete Assembly	SRA: CP011147.1 [16]
MRSA1	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	SRA: BX571857-8 [4]
MRSA2	United Kingdom	2011/2012	MRSA	1	1	-	-	Reads	ENA: ERR108055 [6]
MRSA3	United Kingdom	2011/2012	MRSA	1	1	-	-	Reads	ENA: ERR108057 [6]
MRSA4	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	ENA: ERR108058 [6]
MRSA5	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	ENA: ERR108059 [6]
MRSA7	United Kingdom	2011/2012	MRSA	5	5	-	-	Reads	ENA: ERR108061 [6]
MRSA10	United Kingdom	2011/2012	MRSA	5	5	-	-	Reads	ENA: ERR108040 [6]
MRSA13	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	ENA: ERR108043 [6]
MRSA15	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	ENA: ERR108045 [6]
MRSA16	United Kingdom	2011/2012	MRSA	45	45	-	-	Reads	ENA: ERR108046 [6]
MRSA17	United Kingdom	2011/2012	MRSA	149	5	-	-	Reads	ENA: ERR108047 [6]
MRSA18	United Kingdom	2011/2012	MRSA	8	8	-	-	Reads	ENA: ERR108048 [6]
Sa_NZ1001	New Zealand	2013	MRSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ1002	New Zealand	2013	MSSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ1003	New Zealand	2013	MSSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ1004	New Zealand	2013	MSSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ1005	New Zealand	2013	MSSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ1006	New Zealand	2013	MSSA	1	1	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZAK3	New Zealand	2005	MRSA	5	5	Resistant	PacBio RS-II	Complete Assembly	This Study
Sa_NZ5002	New Zealand	2006	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5003	New Zealand	2006	MRSA	149	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5004	New Zealand	2006	MRSA	73	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5005	New Zealand	2006	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5006	New Zealand	2006	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5007	New Zealand	2007	MRSA	149	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5008	New Zealand	2007	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5009	New Zealand	2007	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5010	New Zealand	2007	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5011	New Zealand	2008	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5012	New Zealand	2008	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5013	New Zealand	2008	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5014	New Zealand	2009	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5015	New Zealand	2009	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5016	New Zealand	2009	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5017	New Zealand	2009	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5018	New Zealand	2009	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5019	New Zealand	2009	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5020	New Zealand	2010	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5021	New Zealand	2010	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5022	New Zealand	2010	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5023	New Zealand	2010	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5024	New Zealand	2011	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5025	New Zealand	2011	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5026	New Zealand	2011	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5027	New Zealand	2011	MRSA	5	5	Susceptible	Illumina MiSeq	Reads	This Study
Sa_NZ5028	New Zealand	2011	MSSA	149	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5029	New Zealand	2011	MRSA	5	5	Resistant	Illumina MiSeq	Reads	This Study
Sa_NZ5030	New Zealand	2012	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5031	New Zealand	2012	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5032	New Zealand	2012	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5033	New Zealand	2013	MSSA	835	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5034	New Zealand	2013	MSSA	835	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5035	New Zealand	2013	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5036	New Zealand	2013	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5037	New Zealand	2013	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study
Sa_NZ5038	New Zealand	2013	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5039	New Zealand	2013	MRSA	5	5	Resistant	Illumina HiSeq	Reads	This Study
Sa_NZ5040	New Zealand	2013	MSSA	5	5	Susceptible	Illumina HiSeq	Reads	This Study

Abbreviations: MLST, multilocus sequence type; MSSA, methicillin-susceptible *Staphylococcus aureus*; MRSA, methicillin-resistant *S. aureus*; WGS, whole genome sequence; SRA, short read archive; ENA, European Nucleotide Archive* MLST of New Zealand isolates was determined *in silico*

Supplementary Dataset

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
1	NAZK3	-	2005	New Zealand	Complete Genomes	PacBio RS-II	LT009690,LT009691	-	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
2	502A	-	-	USA	Complete Genomes	PacBio RS-II	CP007454, CP007455	-	Parker (2014) PMID: 24723721	Parker (2014) PMID: 24723721
3	ECT-R2	-	2004-07	Sweden	Complete Genomes	454 Pyro-Sequencing	FR714927, FR714928, FR714929	-	Lindqvist (2012) PMID: 21590357	Lindqvist (2012) PMID: 21590357
4	ED98	-	1996-97	Ireland	Complete Genomes	454 Pyro-Sequencing	CP001781, CP001782, CP001783, CP001784	-	Lowder (2009) PMID: 19884497	Lowder (2009) PMID: 19884497
5	HOU1444-VR	-	-	-	Complete Genomes	PacBio RS-II	CP012593, CP012594, CP012595, CP012596	-	Panesso (2015) PMID: 26402569	Panesso (2015) PMID: 26402569
6	JH1	-	2000	USA	Complete Genomes	-	CP000736, CP000737	-	Mwangi (2007) PMID: 17517606	Mwangi (2007) PMID: 17517606
7	JH9	-	2000	USA	Complete Genomes	-	CP000703, CP000704	-	Mwangi (2007) PMID: 17517606	Mwangi (2007) PMID: 17517606
8	Mu3	-	1996	Japan	Complete Genomes	-	AP009324	-	Neoh (2008) PMID: 17954695	Neoh (2008) PMID: 17954695
9	Mu50	-	1996	Japan	Complete Genomes	-	BA000017, AP003367	-	Kuroda (2001) PMID: 11418146	Kuroda (2001) PMID: 11418146
10	N315	-	1982	Japan	Complete Genomes	-	BA000018, AP003139	-	Kuroda (2001) PMID: 11418146	Kuroda (2001) PMID: 11418146
11	SA564	-	-	-	Complete Genomes	Illumina HiSeq 2500	CP010890, CP010891	-	Giraud (2015) PMID: 25997461	Giraud (2015) PMID: 25997461
12	FCFHV36	-	-	Brazil	Complete Genomes	Illumina MiSeq	CP011147	-	McCulloch (2015) PMID: 26272570	McCulloch (2015) PMID: 26272570
13	04-02981	-	2004	Germany	Complete Genomes	454 GS FLX Titanium	CP001844	-	Nubel (2010) PMID: 20386717	Nubel (2010) PMID: 20386717
14	H10388	I1	2001	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579059, HE579060	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
15	H10497	I2	2001	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579061, HE579062	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
16	H15532	I3	2006	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579063, HE579064	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
17	H16035	I4	2006	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579065, HE579066	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
18	H16125	I5	2006	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579067, HE579068	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
19	H18341	I6	2008	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579069, HE579070	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
20	H18412	I7	2008	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579071, HE579072	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
21	H18583	I8	2008	Switzerland	Complete Genomes	Illumina Genome Analyzer Ix	HE579073, HE579074	-	Vogel (2012) PMID: 22720005	Vogel (2012) PMID: 22720005
22	NZ5001	NZAK3	2005	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723493	ERR1213699	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
23	NZ5002	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723500	ERR1201317	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
24	NZ5003	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723501	ERR1201318	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
25	NZ5004	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723502	ERR1201319	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
26	NZ5005	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723503	ERR1201320	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
27	NZ5006	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723504	ERR1201321	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
28	NZ5007	-	2007	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723505	ERR1201322	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
29	NZ5008	-	2007	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723506	ERR1201323	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
30	NZ5009	-	2007	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723507	ERR1201324	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
31	NZ5010	-	2007	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723508	ERR1201325	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
32	NZ5011	-	2008	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723509	ERR1201326	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
33	NZ5012	-	2008	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723510	ERR1201327	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
34	NZ5013	-	2008	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723511	ERR1201328	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
35	NZ5014	-	2009	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723512	ERR1201329	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
36	NZ5015	-	2009	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723513	ERR1201330	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
37	NZ5016	-	2009	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723514	ERR1201331	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
38	NZ5017	-	2009	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723515	ERR1201332	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
39	NZ5018	-	2009	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723516	ERR1201333	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
40	NZ5019	-	2009	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723517	ERR1201334	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
41	NZ5020	-	2010	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723518	ERR1201335	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
42	NZ5021	-	2010	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723519	ERR1201336	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
43	NZ5022	-	2010	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723520	ERR1201337	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
44	NZ5023	-	2010	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723521	ERR1201338	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
45	NZ5024	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723522	ERR1201339	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
46	NZ5025	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723523	ERR1201340	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
47	NZ5026	-	2011	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723524	ERR1201341	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
48	NZ5027	-	2011	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723525	ERR1201342	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
49	NZ5028	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723526	ERR1201343	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
50	NZ5029	-	2011	New Zealand	WGS Reads	Illumina MiSeq	SAMEA3723527	ERR1201344	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
51	NZ5030	-	2012	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723528	ERR1201345	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
52	NZ5031	-	2012	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723529	ERR1201346	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
53	NZ5032	-	2012	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723530	ERR1201347	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
54	NZ5033	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723531	ERR1201348	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
55	NZ5034	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723532	ERR1201349	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
56	NZ5035	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723533	ERR1201350	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
57	NZ5036	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723534	ERR1201351	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
58	NZ5037	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723535	ERR1201352	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
59	NZ5038	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723536	ERR1201353	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
60	NZ5039	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723537	ERR1201354	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
61	NZ5040	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	SAMEA3723538	ERR1201355	Baines (2016) PMID: 26856837	Baines (2016) PMID: 26856837
62	NZ5042	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
63	NZ5043	-	2006	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
64	NZ5048	-	2008	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
65	NZ5049	-	2008	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
66	NZ5050	-	2008	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
67	NZ5051	-	2009	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
68	NZ5052	-	2010	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
69	NZ5053	-	2010	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
70	NZ5054	-	2010	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
71	NZ5055	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
72	NZ5056	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
73	NZ5057	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
74	NZ5058	-	2012	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
75	NZ5059	-	2011	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
76	NZ5061	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
77	NZ5062	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
78	NZ5063	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
79	NZ5064	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
80	NZ5065	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
81	NZ5066	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
82	NZ5067	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
83	NZ5068	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
84	NZ5069	-	2013	New Zealand	WGS Reads	Illumina HiSeq 2000	-	-	Unpublished	Unpublished
85	MRSA1	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317989	ERR108055	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
86	MRSA4	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317994	ERR108058	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
87	MRSA5	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317977	ERR108059	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
88	MRSA7	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317988	ERR108061	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
89	MRSA10	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317982	ERR108040	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
90	MRSA13	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317981	ERR108043	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
91	MRSA15	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317983	ERR108045	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
92	MRSA17	-	2011/12	United Kingdom	WGS Reads	Illumina MiSeq	SAMEA1317996	ERR108047	Ellington (2015) PMID: 25769787	Ellington (2015) PMID: 25769787
93	BPH2819	-	2011	Australia	WGS Reads	Illumina HiSeq 2000	-	-	Holmes (2018) PMID: 29506483	Unpublished
94	BPH2715	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
95	BPH2716	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
96	BPH2719	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
97	BPH2743	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
98	BPH2754	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
99	BPH2757	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
100	BPH2758	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
101	BPH2786	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
102	BPH2793	-	2011	Australia	WGS Reads	Illumina HiSeq 2000	-	-	Holmes (2018) PMID: 29506483	Unpublished
103	BPH2794	-	2011	Australia	WGS Reads	Illumina HiSeq 2000	-	-	Holmes (2018) PMID: 29506483	Unpublished
104	BPH2818	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
105	BPH2883	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
106	BPH2885	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
107	BPH2886	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
108	BPH2901	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
109	BPH2913	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
110	BPH2914	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
111	BPH2933	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
112	BPH2936	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
113	BPH2939	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
114	BPH2942	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
115	BPH2955	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
116	BPH2959	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
117	BPH2985	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
118	BPH2881	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
119	BPH2809	-	2011	Australia	WGS Reads	Illumina HiSeq 2000	-	-	Holmes (2018) PMID: 29506483	Unpublished
120	BPH2767	-	2011	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
121	BPH2849	-	2012	Australia	WGS Reads	Illumina MiSeq	-	-	Holmes (2018) PMID: 29506483	Unpublished
122	BPH2872	-	2012	Australia	WGS Reads	Illumina HiSeq 2000	-	-	Holmes (2018) PMID: 29506483	Unpublished
123	BPH3214	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
124	BPH3217	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
125	BPH3227	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
126	BPH3232	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
127	BPH3240	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
128	BPH3241	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
129	BPH3252	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
130	BPH3263	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
131	BPH3273	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
132	BPH3274	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
133	BPH3282	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
134	BPH3295	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
135	BPH3309	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
136	BPH3317	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
137	BPH3318	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
138	BPH3321	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
139	BPH3322	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
140	BPH3323	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
141	BPH3329	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
142	BPH3334	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
143	BPH3336	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
144	BPH3347	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
145	BPH3359	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
146	BPH3388	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
147	BPH3391	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
148	BPH3412	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
149	BPH3416	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
150	BPH3419	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
151	BPH3422	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
152	BPH3424	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
153	BPH3428	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
154	BPH3446	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
155	BPH3459	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
156	BPH3462	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
157	BPH3467	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
158	BPH3475	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
159	BPH3480	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
160	BPH3513	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
161	BPH3542	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
162	BPH3556	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
163	BPH3571	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
164	BPH3575	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
165	BPH3579	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
166	BPH3585	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
167	BPH3607	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
168	BPH3610	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
169	BPH3611	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
170	BPH3627	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
171	BPH3635	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
172	BPH3653	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
173	BPH3656	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
174	BPH3659	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
175	BPH3679	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
176	BPH3701	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
177	BPH3716	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
178	BPH3721	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
179	BPH3728	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
180	BPH3582	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
181	BPH3616	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
182	BPH3620	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
183	BPH3530	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
184	BPH3591	-	2007	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
185	BPH3675	-	2008	Australia	WGS Reads	Illumina NextSeq 500	-	-	Holmes (2011) PMID: 21742831	Unpublished
186	S20106-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
187	S20111-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
188	S20112-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
189	S20302-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
190	S20405-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
191	S20614-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
192	S20807-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
193	S20902-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
194	S31914-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
195	S31920-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
196	S32101-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
197	S32102-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
198	S32302-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
199	S41105-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
200	S41109-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
201	S41114-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
202	S41205-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
203	S41211-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
204	S41216-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
205	S41301-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
206	S42902-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
207	S42914-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
208	S43001-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
209	S51002-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
210	S51015-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
211	S51405-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
212	S51407-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
213	S51409-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
214	S51501-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
215	S51512-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
216	S51517-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
217	S51530-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
218	S51602-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
219	S51608-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
220	S51609-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
221	S53504-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
222	S53506-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
223	S62501-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
224	S62610-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
225	S63301-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
226	S63302-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
227	S63303-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
228	S63708-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
229	S63709-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
230	S63710-2015	-	2015	Australia	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys	Unpublished
231	EUH15	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649751	SRR1177884	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
232	4597	2006734597	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649743	SRR1177876	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
233	5535	2006745535	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649744	SRR1177877	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
234	8535	2003718535	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649747	SRR1177880	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
235	GA48963	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649754	SRR1177887	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
236	GA50819	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649756	SRR1177889	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
237	CA598	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649775	SRR1177908	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
238	NY593	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649789	SRR1177922	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
239	CA26	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649748	SRR1177881	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
240	CA541	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649749	SRR1177882	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
241	GA53649	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649759	SRR1177892	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
242	SR4155	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649766	SRR1177899	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
243	SR4156	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649767	SRR1177900	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
244	CA11	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649770	SRR1177903	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
245	CA39	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649771	SRR1177904	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
246	CA436	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649772	SRR1177905	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
247	CA544	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649773	SRR1177906	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
248	CA545	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649774	SRR1177907	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
249	CA9	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649776	SRR1177909	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
250	GA12	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649777	SRR1177910	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
251	GA984	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649780	SRR1177913	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
252	MN055	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649781	SRR1177914	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
253	NY341	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649784	SRR1177917	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
254	NY439	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649788	SRR1177921	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
255	SR3569	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649794	SRR1177927	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
256	SR3777	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649796	SRR1177929	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
257	SR4153	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649798	SRR1177931	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
258	SR4187	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649799	SRR1177932	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
259	SR3585	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649764	SRR1177897	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
260	SR4152	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649765	SRR1177898	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
261	MN105a	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649782	SRR1177915	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
262	NY224	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649783	SRR1177916	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
263	NY356	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649785	SRR1177918	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
264	NY360	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649786	SRR1177919	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
265	NY417	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649787	SRR1177920	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
266	SR220b	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649793	SRR1177926	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
267	SR3732	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649795	SRR1177928	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
268	407a	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649811	SRR1177944	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
269	107	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649802	SRR1177935	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
270	115c	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649804	SRR1177937	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
271	122VISA	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649805	SRR1177938	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
272	217	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649808	SRR1177941	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
273	221	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649809	SRR1177942	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
274	506	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649813	SRR1177946	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
275	123a	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649806	SRR1177939	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
276	504a	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649812	SRR1177945	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
277	120336	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649800	SRR1177933	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
278	120335	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649768	SRR1177901	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
279	120337	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649801	SRR1177934	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
280	120338	-	2005	United States	WGS Reads	Illumina HiSeq 2000	SAMN02649815	SRR1177948	Alam (2014) PMID:24787619	Alam (2014) PMID:24787619
281	US_086	-	2012	Thailand	WGS Reads	Illumina HiSeq 2000	AMEA3449000	ERR1069853	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
282	US_095	-	2012	Thailand	WGS Reads	Illumina HiSeq 2000	AMEA3449005	ERR1069858	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
283	US_100	-	2012	Thailand	WGS Reads	Illumina HiSeq 2000	AMEA3449010	ERR1069863	Chantratita (2016) PMID: 26806258	Moradigaravand (2017) PMID: 28679748
284	US_103	-	2012	Thailand	WGS Reads	Illumina HiSeq 2000	AMEA3449012	ERR1069865	Chantratita (2016) PMID: 26806258	Moradigar

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
320	WCH0158	-	2010	United States	WGS Reads	Illumina MiSeq	SAMN05728320	SRR4140034	Azarian (2016) PMID:27732618	Azarian (2016) PMID:27732618
321	WCH0168	-	2010	United States	WGS Reads	Illumina MiSeq	SAMN05728321	SRR4140035	Azarian (2016) PMID:27732618	Azarian (2016) PMID:27732618
322	WCH0183	-	2004	United States	WGS Reads	Illumina MiSeq	SAMN05728322	SRR4140036	Azarian (2016) PMID:27732618	Azarian (2016) PMID:27732618
323	WCH0186	-	2004	United States	WGS Reads	Illumina MiSeq	SAMN05728323	SRR4140037	Azarian (2016) PMID:27732618	Azarian (2016) PMID:27732618
324	GR07A	-	2012	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536974	ERR599944	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
325	GUH23	-	2012	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536935	ERR599956	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
326	GR13A	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536984	ERR599771	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
327	GR14A	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536985	ERR599959	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
328	GUH34A	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536950	ERR599921	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
329	GR14B	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536997	ERR599960	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
330	GR13B	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2536999	ERR599782	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
331	GR14C	-	2013	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA2537000	ERR599962	Harrison (2016) PMID:27716432	Harrison (2016) PMID:27716432
332	87	A3	2009	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727081	ERR1213811	Senghore (2016) PMID:27474712	Senghore (2016) PMID:27474712
333	135	A5	2002	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727091	ERR1213821	Senghore (2016) PMID:27474712	Senghore (2016) PMID:27474712
334	860	SA17	2005	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727060	ERR1213790	Senghore (2016) PMID:27474712	Senghore (2016) PMID:27474712
335	98128	SA16	2004	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727052	ERR1213782	Ebruke (2016) PMID: 26969294	Ebruke (2016) PMID: 26969294
336	116431	SA3	2004	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727054	ERR1213784	Ebruke (2016) PMID: 26969294	Ebruke (2016) PMID: 26969294
337	104528	SA14	2004	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727050	ERR1213780	Ebruke (2016) PMID: 26969294	Ebruke (2016) PMID: 26969294
338	103277	SA15	2004	Gambia	WGS Reads	Illumina MiSeq	SAMEA3727051	ERR1213781	Ebruke (2016) PMID: 26969294	Ebruke (2016) PMID: 26969294
339	BSAC_st45	-	2001	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1317462	ERR107800	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
340	BSAC_st48	-	2001	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1317583	ERR107801	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
341	BSAC_st527	-	2003	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1317575	ERR107892	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
342	BSAC_st970	-	2004	England	WGS Reads	Illumina HiSeq 2000	SAMEA1317493	ERR107971	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
343	BSAC_st1792	-	2008	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1317469	ERR111046	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
344	BSAC_st1823	-	2008	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA1317453	ERR111058	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
345	BSAC_st1827	-	2008	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA1317405	ERR111061	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
346	BSAC_st1193	-	2005	England	WGS Reads	Illumina HiSeq 2000	SAMEA1469669	ERR158648	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
347	BSAC_st1307	-	2006	England	WGS Reads	Illumina HiSeq 2000	SAMEA1469581	ERR158691	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
348	BSAC_st2293	-	2009	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1469613	ERR158744	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
349	BSAC_st1062	-	2005	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464639	ERR118363	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
350	BSAC_st1664	-	2007	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464719	ERR118490	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
351	BSAC_st124	-	2001	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464181	ERR118529	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
352	BSAC_st228	-	2001	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464466	ERR118568	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
353	BSAC_st277	-	2002	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464600	ERR118586	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
354	BSAC_st774	-	2004	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464454	ERR129305	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
355	BSAC_st776	-	2004	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464217	ERR129306	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
356	BSAC_st659	-	2003	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464480	ERR129259	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
357	BSAC_st1815	-	2008	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA1464546	ERR134357	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
358	BSAC_st326	-	2002	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464437	ERR134334	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
359	BSAC_st2661	-	2009	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464475	ERR134384	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
360	BSAC_st2801	-	2010	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464381	ERR134410	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
361	BSAC_st2805	-	2010	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464542	ERR134412	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
362	BSAC_st586	-	2003	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464299	ERR134340	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
363	BSAC_st852	-	2004	Ireland	WGS Reads	Illumina HiSeq 2000	SAMEA1464464	ERR129332	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
364	BSAC_st2761	-	2010	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464721	ERR134399	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
365	BSAC_st1028	-	2005	Wales	WGS Reads	Illumina HiSeq 2000	SAMEA1464230	ERR118356	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
366	BSAC_st1424	-	2006	England	WGS Reads	Illumina HiSeq 2000	SAMEA1464594	ERR134348	Reuter (2016) PMID:26672018	Reuter (2016) PMID:26672018
367	754	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743501	ERR593824	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
368	782	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743744	ERR594068	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
369	968	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743845	ERR594162	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
370	1055	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2742908	ERR593111	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
371	1063	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743741	ERR594065	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
372	1354	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743803	ERR594130	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
373	1358	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743822	ERR594150	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
374	2632	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716941	ERR1195900	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
375	2642	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716953	ERR1195913	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
376	2800	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716989	ERR1195956	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
377	2804	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716994	ERR1195961	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
378	2808	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716996	ERR1195965	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
379	2858	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3717028	ERR1196002	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
380	283	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2742910	ERR593113	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
381	427	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2765399	ERR598731	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
382	590	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2743035	ERR593236	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
383	724	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743288	ERR593610	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178

	ID	Alternate ID	Year	Country	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Isolate Reference	Sequence Reference
384	833	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743782	ERR594105	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
385	1367	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743819	ERR594147	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
386	2499	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA3717882	ERR1197500	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
387	2625	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716934	ERR1195890	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
388	2919	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3717128	ERR1197193	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
389	3011	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3717177	ERR1197248	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
390	2788	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716986	ERR1195953	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
391	2941	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3717130	ERR1197195	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
392	2774	-	2015	Germany	WGS Reads	Illumina MiSeq	SAMEA3716977	ERR1195942	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
393	2448	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA3716839	ERR1195779	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
394	1023	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2746598	ERR594218	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
395	1335	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743790	ERR594115	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
396	485	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2742967	ERR593163	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
397	965	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2744118	ERR594172	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
398	2574	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA3716897	ERR1195844	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
399	505	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2742983	ERR593179	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
400	1344	-	2014	Germany	WGS Reads	Illumina MiSeq	SAMEA2743807	ERR594134	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
401	424	-	2013	Germany	WGS Reads	Illumina MiSeq	SAMEA2742929	ERR593132	Mellmann (2016) PMID:27558178	Mellmann (2016) PMID:27558178
402	Sa_NOG-W33	Sa_NOG-W33	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
403	Sa_NOG-W35	Sa_NOG-W35	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
404	Sa_NOG-W36	Sa_NOG-W36	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
405	Sa_NOG-W64	Sa_NOG-W64	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
406	Sa_NOG-W71	Sa_NOG-W71	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
407	Sa_NOG-W72	Sa_NOG-W72	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Kpeli (2016) PMID:27547406	Kpeli (2016) PMID:27547406
408	N34_260215	N34_260215	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Amissah (2017) PMID: 29056927	Amissah (2017) PMID: 29056927
409	W31_250215	W31_250215	-	Ghana	WGS Reads	Illumina HiSeq 2000	-	-	Amissah (2017) PMID: 29056927	Amissah (2017) PMID: 29056927

Appendix 3

Chapter 4: Supplementary Material

Publication: Supplementary Dataset S1

Supplementary Dataset

	ID	Year	Country	State (AUS)	Data Type	Sequence Platform	Sample/Genome Accession	Run Accession	Sequence Reference
1	BPH0468	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212938	ERR732815	Baines. 2015. mBio. PMID: 25736880
2	BPH0475	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212931	ERR732816	Baines. 2015. mBio. PMID: 25736880
3	BPH0488	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212956	ERR732817	Baines. 2015. mBio. PMID: 25736880
4	BPH0491	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212955	ERR732818	Baines. 2015. mBio. PMID: 25736880
5	BPH0494	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212940	ERR732819	Baines. 2015. mBio. PMID: 25736880
6	BPH0496	2004	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212928	ERR732820	Baines. 2015. mBio. PMID: 25736880
7	BPH0499	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212937	ERR732821	Baines. 2015. mBio. PMID: 25736880
8	BPH0502	2003	Australia	WA	WGS Reads	Illumina MiSeq	SAMEA3212919	ERR732822	Baines. 2015. mBio. PMID: 25736880
9	BPH0503	2003	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212967	ERR732823	Baines. 2015. mBio. PMID: 25736880
10	BPH0508	2006	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212893	ERR732824	Baines. 2015. mBio. PMID: 25736880
11	BPH0509	2005	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212911	ERR732825	Baines. 2015. mBio. PMID: 25736880
12	BPH0510	2005	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212957	ERR732826	Baines. 2015. mBio. PMID: 25736880
13	BPH2002	2002	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212904	ERR732827	Baines. 2015. mBio. PMID: 25736880
14	BPH2003	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212913	ERR732828	Baines. 2015. mBio. PMID: 25736880
-	pBPH2003 (BPH2003)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607265	-	This Study
-	pBPH2003 (BPH2003)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA5047398	ERR2858549	This Study
15	BPH2004	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212949	ERR732829	Baines. 2015. mBio. PMID: 25736880
16	BPH2005	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212961	ERR732830	Baines. 2015. mBio. PMID: 25736880
17	BPH2006	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212925	ERR732831	Baines. 2015. mBio. PMID: 25736880
18	BPH2007	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212952	ERR732832	Baines. 2015. mBio. PMID: 25736880
19	BPH2008	2000	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212918	ERR732833	Baines. 2015. mBio. PMID: 25736880
20	BPH2009	2000	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212948	ERR732834	Baines. 2015. mBio. PMID: 25736880
21	BPH2010	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212899	ERR732835	Baines. 2015. mBio. PMID: 25736880
22	BPH2011	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212962	ERR732836	Baines. 2015. mBio. PMID: 25736880
23	BPH2012	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212958	ERR732837	Baines. 2015. mBio. PMID: 25736880
24	BPH2013	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212934	ERR732838	Baines. 2015. mBio. PMID: 25736880
25	BPH2014	1999	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212969	ERR732839	Baines. 2015. mBio. PMID: 25736880
26	BPH2015	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212954	ERR732840	Baines. 2015. mBio. PMID: 25736880
27	BPH2016	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212908	ERR732841	Baines. 2015. mBio. PMID: 25736880
28	BPH2017	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212936	ERR732842	Baines. 2015. mBio. PMID: 25736880
29	BPH2018	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212942	ERR732843	Baines. 2015. mBio. PMID: 25736880
30	BPH2019	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212951	ERR732844	Baines. 2015. mBio. PMID: 25736880
-	SV3 (BPH2019)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607255	-	This Study
-	SV3 (BPH2019)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA5047399	ERR2858550	This Study
31	BPH2021	2004	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212903	ERR732845	Baines. 2015. mBio. PMID: 25736880
32	BPH2022	2005	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212945	ERR732846	Baines. 2015. mBio. PMID: 25736880
33	BPH2023	2006	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212915	ERR732847	Baines. 2015. mBio. PMID: 25736880
34	BPH2024	2007	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212944	ERR732848	Baines. 2015. mBio. PMID: 25736880
35	BPH2029	2009	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212905	ERR732852	Baines. 2015. mBio. PMID: 25736880
36	BPH2031	2008	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212965	ERR732853	Baines. 2015. mBio. PMID: 25736880
37	BPH2032	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212926	ERR732854	Baines. 2015. mBio. PMID: 25736880
38	BPH2034	2008	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212917	ERR732856	Baines. 2015. mBio. PMID: 25736880
39	BPH2054	1980	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212883	ERR732860	Baines. 2015. mBio. PMID: 25736880
40	BPH2055	1980	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212885	ERR732861	Baines. 2015. mBio. PMID: 25736880
41	BPH2056	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212902	ERR732862	Baines. 2015. mBio. PMID: 25736880
-	SV1 (BPH2056)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607305	-	This Study
-	SV1 (BPH2056)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA5047400	ERR2858551	This Study
42	BPH2057	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212889	ERR732863	Baines. 2015. mBio. PMID: 25736880

43	BPH2059	1981	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212897	ERR732864	Baines. 2015. mBio. PMID: 25736880
44	BPH2068	2009	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212906	ERR732865	Baines. 2015. mBio. PMID: 25736880
45	BPH2070	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212890	ERR732866	Baines. 2015. mBio. PMID: 25736880
-	SV2 (BPH2070)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607275	-	This Study
-	SV2 (BPH2070)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA5047401	ERR2858552	This Study
46	BPH2071	1998	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212888	ERR732867	Baines. 2015. mBio. PMID: 25736880
47	BPH2072	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212947	ERR732868	Baines. 2015. mBio. PMID: 25736880
48	BPH2073	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212927	ERR732869	Baines. 2015. mBio. PMID: 25736880
49	BPH2075	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212896	ERR732870	Baines. 2015. mBio. PMID: 25736880
50	BPH2077	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212932	ERR732871	Baines. 2015. mBio. PMID: 25736880
51	BPH2078	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212950	ERR732872	Baines. 2015. mBio. PMID: 25736880
52	BPH2079	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212922	ERR732873	Baines. 2015. mBio. PMID: 25736880
53	BPH2080	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212929	ERR732874	Baines. 2015. mBio. PMID: 25736880
54	BPH2081	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212941	ERR732875	Baines. 2015. mBio. PMID: 25736880
55	BPH2082	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212920	ERR732876	Baines. 2015. mBio. PMID: 25736880
56	BPH2083	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212886	ERR732877	Baines. 2015. mBio. PMID: 25736880
57	BPH2084	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212892	ERR732878	Baines. 2015. mBio. PMID: 25736880
58	BPH2085	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212907	ERR732879	Baines. 2015. mBio. PMID: 25736880
59	BPH2086	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212900	ERR732880	Baines. 2015. mBio. PMID: 25736880
60	BPH2088	1982	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA3212968	ERR732881	Baines. 2015. mBio. PMID: 25736880
61	BPH2090	2012	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212943	ERR732882	Baines. 2015. mBio. PMID: 25736880
62	BPH2094	1998	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212946	ERR732883	Baines. 2015. mBio. PMID: 25736880
63	BPH2095	2003	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212887	ERR732884	Baines. 2015. mBio. PMID: 25736880
64	BPH2096	2005	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212923	ERR732885	Baines. 2015. mBio. PMID: 25736880
65	BPH2097	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212939	ERR732886	Baines. 2015. mBio. PMID: 25736880
66	BPH2098	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212884	ERR732887	Baines. 2015. mBio. PMID: 25736880
67	BPH2099	2007	Australia	NSW	WGS Reads	Ion Torrent PGM	SAMEA3212933	ERR732888	Baines. 2015. mBio. PMID: 25736880
68	JKD6000	2002	Australia	VIC	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212909	ERR732889	Baines. 2015. mBio. PMID: 25736880
69	JKD6004	2002	Australia	QLD	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212901	ERR732891	Baines. 2015. mBio. PMID: 25736880
70	JKD6008	2006	New Zealand	-	Complete Genome	Roache 454 GS20, SOLiD	CP002120	-	Howden. 2010. JBact. PMID:20802046
71	JKD6009	2003	New Zealand	-	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212894	ERR732894	Baines. 2015. mBio. PMID: 25736880
-	SV5 (JKD6009)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607245	-	This Study
-	SV5 (JKD6009)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA5047402	ERR2858553	This Study
72	JKD6021	2001	Australia	VIC	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212953	ERR732895	Baines. 2015. mBio. PMID: 25736880
73	JKD6052	2004	Australia	QLD	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212916	ERR732899	Baines. 2015. mBio. PMID: 25736880
74	JKD6121	2005	Australia	VIC	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212963	ERR732900	Baines. 2015. mBio. PMID: 25736880
75	BPH2065	1981	Australia	SA	WGS Reads	Illumina NextSeq 500	SAMEA5042664	ERR2855733	This Study
76	BPH2066	1981	Australia	SA	WGS Reads	Illumina NextSeq 500	SAMEA5042665	ERR2855734	This Study
77	BPH2104	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042666	ERR2855735	This Study
78	BPH2106	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042667	ERR2855736	This Study
79	BPH2107	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042669	ERR2855737	This Study
80	BPH2154	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042676	ERR2855744	This Study
81	BPH2161	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042678	ERR2855746	This Study
82	BPH2163	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042679	ERR2855747	This Study
83	BPH2165	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042680	ERR2855748	This Study
84	BPH2203	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042687	ERR2855755	This Study
85	BPH2205	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042689	ERR2855757	This Study
86	BPH2226	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042695	ERR2855763	This Study
87	BPH2228	1990	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA5042696	ERR2855764	This Study
88	BPH2233	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042699	ERR2855767	This Study
89	BPH2234	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042700	ERR2855768	This Study

90	BPH2235	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042701	ERR2855769	This Study
91	BPH2236	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042702	ERR2855770	This Study
92	BPH2238	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042703	ERR2855771	This Study
93	BPH2710	2011	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317322	ERR2137022	Publication in process
94	BPH2718	2011	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317330	ERR2137030	Publication in process
95	BPH2723	2011	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317335	ERR2137035	Publication in process
96	BPH2726	2011	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317338	ERR2137038	Publication in process
97	BPH2747	2011	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317359	ERR2137059	Publication in process
98	BPH2779	2011	Australia	VIC	WGS Reads	Illumina HiSeq	SAMEA104317390	ERR2137090	Publication in process
99	BPH2780	2011	Australia	VIC	WGS Reads	Illumina HiSeq	SAMEA104317391	ERR2137091	Publication in process
100	BPH2781	2011	Australia	VIC	WGS Reads	Illumina HiSeq	SAMEA104317392	ERR2137092	Publication in process
101	BPH2787	2011	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA104317398	ERR2137098	Publication in process
102	BPH2798	2011	Australia	VIC	WGS Reads	Illumina HiSeq	SAMEA104317409	ERR2137109	Publication in process
103	BPH2824	2011	Australia	VIC	WGS Reads	Illumina HiSeq	SAMEA104317431	ERR2137131	Publication in process
104	BPH2827	2011	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA104317434	ERR2137134	Publication in process
105	BPH2869	2012	Australia	NSW	WGS Reads	Illumina MiSeq	SAMEA104317472	ERR2137172	Publication in process
-	SV6 (BPH2869)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607295	-	This Study
-	SV6 (BPH2869)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA104317472	ERR2858554	This Study
106	BPH2896	2012	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317497	ERR2137197	Publication in process
107	BPH2946	2012	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317543	ERR2137243	Publication in process
108	BPH2947	2012	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA104317544	ERR2137244	Publication in process
109	BPH3202	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317594	ERR2137294	Publication in process
110	BPH3208	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317600	ERR2137300	Publication in process
111	BPH3220	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317612	ERR2137312	Publication in process
112	BPH3234	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317626	ERR2137326	Publication in process
113	BPH3244	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317636	ERR2137336	Publication in process
-	SV4 (BPH3244)	-	-	-	Complete Genome	Pacific Biosciences RS-II	GCA_900607285	-	This Study
-	SV4 (BPH3244)	-	-	-	WGS Long Reads	Pacific Biosciences RS-II	SAMEA104317636	ERR2858555	This Study
114	BPH3250	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317642	ERR2137342	Publication in process
115	BPH3254	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317646	ERR2137346	Publication in process
116	BPH3260	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317652	ERR2137352	Publication in process
117	BPH3262	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317654	ERR2137354	Publication in process
118	BPH3266	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317658	ERR2137358	Publication in process
119	BPH3272	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317664	ERR2137364	Publication in process
120	BPH3276	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317668	ERR2137368	Publication in process
121	BPH3284	2008	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317676	ERR2137376	Publication in process
122	BPH3306	2007	Australia	QLD	WGS Reads	Illumina NextSeq 500	SAMEA104317698	ERR2137398	Publication in process
123	BPH3324	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317716	ERR2137416	Publication in process
124	BPH3328	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317720	ERR2137420	Publication in process
125	BPH3332	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317724	ERR2137424	Publication in process
126	BPH3344	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317736	ERR2137436	Publication in process
127	BPH3346	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317738	ERR2137438	Publication in process
128	BPH3348	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317740	ERR2137440	Publication in process
129	BPH3356	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317748	ERR2137448	Publication in process
130	BPH3360	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317752	ERR2137452	Publication in process
131	BPH3362	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317754	ERR2137454	Publication in process
132	BPH3366	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317758	ERR2137458	Publication in process
133	BPH3368	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317760	ERR2137460	Publication in process
134	BPH3370	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317762	ERR2137462	Publication in process
135	BPH3372	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317764	ERR2137464	Publication in process
136	BPH3376	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317768	ERR2137468	Publication in process

137	BPH3382	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317774	ERR2137474	Publication in process
138	BPH3394	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317784	ERR2137484	Publication in process
139	BPH3401	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317791	ERR2137491	Publication in process
140	BPH3403	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317793	ERR2137493	Publication in process
141	BPH3413	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317802	ERR2137502	Publication in process
142	BPH3415	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317804	ERR2137504	Publication in process
143	BPH3417	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317806	ERR2137506	Publication in process
144	BPH3431	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317820	ERR2137520	Publication in process
145	BPH3433	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317822	ERR2137522	Publication in process
146	BPH3435	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317824	ERR2137524	Publication in process
147	BPH3439	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317828	ERR2137528	Publication in process
148	BPH3443	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317832	ERR2137532	Publication in process
149	BPH3451	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317839	ERR2137539	Publication in process
150	BPH3453	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317841	ERR2137541	Publication in process
151	BPH3455	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317843	ERR2137543	Publication in process
152	BPH3457	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317845	ERR2137545	Publication in process
153	BPH3463	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317851	ERR2137551	Publication in process
154	BPH3465	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317853	ERR2137553	Publication in process
155	BPH3469	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317857	ERR2137557	Publication in process
156	BPH3483	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317871	ERR2137571	Publication in process
157	BPH3485	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317873	ERR2137573	Publication in process
158	BPH3493	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317881	ERR2137581	Publication in process
159	BPH3495	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317883	ERR2137583	Publication in process
160	BPH3499	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317886	ERR2137586	Publication in process
161	BPH3507	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317894	ERR2137594	Publication in process
162	BPH3511	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317898	ERR2137598	Publication in process
163	BPH3523	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317910	ERR2137610	Publication in process
164	BPH3525	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317912	ERR2137612	Publication in process
165	BPH3527	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317914	ERR2137614	Publication in process
166	BPH3529	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317916	ERR2137616	Publication in process
167	BPH3531	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317918	ERR2137618	Publication in process
168	BPH3533	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317920	ERR2137620	Publication in process
169	BPH3535	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317922	ERR2137622	Publication in process
170	BPH3541	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317928	ERR2137628	Publication in process
171	BPH3543	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317930	ERR2137630	Publication in process
172	BPH3545	2008	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317932	ERR2137632	Publication in process
173	BPH3547	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317934	ERR2137634	Publication in process
174	BPH3549	2007	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA104317936	ERR2137636	Publication in process
175	BPH3632	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325186	ERR2145145	Publication in process

188	BPH3662	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325216	ERR2145175	Publication in process
189	BPH3664	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325218	ERR2145177	Publication in process
190	BPH3666	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325220	ERR2145179	Publication in process
191	BPH3668	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325222	ERR2145181	Publication in process
192	BPH3670	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325224	ERR2145183	Publication in process
193	BPH3672	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325226	ERR2145185	Publication in process
194	BPH3674	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325228	ERR2145187	Publication in process
195	BPH3676	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325230	ERR2145189	Publication in process
196	BPH3678	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325232	ERR2145191	Publication in process
197	BPH3680	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325234	ERR2145193	Publication in process
198	BPH3682	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325236	ERR2145195	Publication in process
199	BPH3684	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325238	ERR2145197	Publication in process
200	BPH3686	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325240	ERR2145199	Publication in process
201	BPH3688	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325242	ERR2145201	Publication in process
202	BPH3692	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325246	ERR2145205	Publication in process
203	BPH3694	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325248	ERR2145207	Publication in process
204	BPH3696	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325250	ERR2145209	Publication in process
205	BPH3698	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325252	ERR2145211	Publication in process
206	BPH3700	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325254	ERR2145213	Publication in process
207	BPH3702	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325256	ERR2145215	Publication in process
208	BPH3704	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325258	ERR2145217	Publication in process
209	BPH3706	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325260	ERR2145219	Publication in process
210	BPH3708	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325262	ERR2145221	Publication in process
211	BPH3710	2008	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325264	ERR2145223	Publication in process
212	BPH3712	2007	Australia	NSW	WGS Reads	Illumina NextSeq 500	SAMEA104325266	ERR2145225	Publication in process

Additional WGS Data

213	JKD6001	2003	Australia	VIC	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212898	ERR732890	Baines. 2015. mBio. PMID: 25736880
214	JKD6005	2002	Australia	QLD	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212895	ERR732892	Baines. 2015. mBio. PMID: 25736880
215	JKD6007	2002	Australia	QLD	WGS Reads	Illumina MiSeq	SAMEA3212912	ERR732893	Baines. 2015. mBio. PMID: 25736880
216	JKD6022	2001	Australia	VIC	WGS Reads	Illumina MiSeq	SAMEA3212914	ERR732896	Baines. 2015. mBio. PMID: 25736880
217	JKD6023	2001	Australia	VIC	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212964	ERR732897	Baines. 2015. mBio. PMID: 25736880
218	JKD6051	2003	Australia	QLD	WGS Reads	Illumina Genome Analyzer IIX	SAMEA3212960	ERR732898	Baines. 2015. mBio. PMID: 25736880
219	BPH2108	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042670	ERR2855738	This Study
220	BPH2109	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042671	ERR2855739	This Study
221	BPH2113	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042672	ERR2855740	This Study
222	BPH2119	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042673	ERR2855741	This Study
223	BPH2120	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042674	ERR2855742	This Study
224	BPH2139	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042675	ERR2855743	This Study
225	BPH2157	1992	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042677	ERR2855745	This Study
226	BPH2167	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042681	ERR2855749	This Study
227	BPH2168	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042682	ERR2855750	This Study
228	BPH2196	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042683	ERR2855751	This Study
229	BPH2198	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042684	ERR2855752	This Study
230	BPH2199	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042685	ERR2855753	This Study
231	BPH2202	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042686	ERR2855754	This Study
232	BPH2204	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042688	ERR2855756	This Study
233	BPH2207	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042690	ERR2855758	This Study
234	BPH2210	1990	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042691	ERR2855759	This Study
235	BPH2213	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042692	ERR2855760	This Study
236	BPH2219	1991	Australia	TAS	WGS Reads	Illumina NextSeq 500	SAMEA5042693	ERR2855761	This Study
237	BPH2222	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042694	ERR2855762	This Study

238	BPH2230	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042697	ERR2855765	This Study
239	BPH2232	1991	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042698	ERR2855766	This Study
240	BPH2240	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042704	ERR2855772	This Study
241	BPH2242	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042705	ERR2855773	This Study
242	BPH2243	1989	Australia	VIC	WGS Reads	Illumina NextSeq 500	SAMEA5042706	ERR2855774	This Study
243	S20209-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
244	S20305-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
245	S20603-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
246	S20607-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
247	S20608-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
248	S20610-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
249	S20613-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
250	S20617-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
251	S20622-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
252	S20623-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
253	S20629-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
254	S20826-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
255	S23616-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
256	S23619-2015	2015	Australia	NSW	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
257	S31902-2015	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
258	S31904-2015	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
259	S31913-2015	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
260	S31918-2015	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
261	S32313-2015	2015	Australia	VIC	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
262	S41210-2015	2015	Australia	QLD	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
263	S51001-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
264	S51003-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
265	S51006-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
266	S51007-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
267	S51008-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
268	S51014-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
269	S51024-2015	2015	Australia	NT	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
270	S51408-2015	2015	Australia	SA	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
271	S51504-2015	2015	Australia	SA	WGS Reads	Illumina MiSeq	-	-	agargroup.org.au/agar-surveys
272	TW20	2003	United Kingdom	-	Complete Genomes	-	FN433596, FN433597, FN433598	-	Holden. 2010. JBact. PMID: 19948800
273	Z172	2010	Taiwan	-	Complete Genomes	Illumina/Solexa + PacBio	CP006838, CP006839, CP006840	-	Chen. 2013. GenomeA. PMCID:PMC3853063
274	T0131	2006	China	-	Complete Genomes	Roche 454 / Solexa	CP002643	-	Li. 2011. JBact. PMID:21551295
275	BMB9393	1993	Brazil	-	Complete Genome	454 GS FLX Titanium	CP005288, CP005289	-	Costa. 2013. GenomeA. PMID:23929475
276	XN108	-	China	-	Complete Genomes	Ion Torrent PGM	CP007447	-	Zhang. 2014. GenomeA. PMID:25059856
277	GV69	1996	Brazil	-	Complete Genomes	454 GS FLX Titanium	CP009681; SAMN03144721	SRR2601051	Botelho. 2016. StandGenSci. PMID:27152133
278	HC1335	2001	Brazil	-	Complete Genomes	454 GS FLX Titanium	CP012012	-	Botelho. 2016. GenBiolEvol. PMID:27635055
279	M92	-	Canada	-	Complete Genomes	PacBio RS-II	CP015447	-	McClure. 2017. GenomeA. PMID:28596402
280	2A8	2001	Czech Republic	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029549	ERR064912	Castillo-Ramirez. 2012. GB. PMID:23270620
281	3HK	2000	Czech Republic	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029566	ERR064907	Castillo-Ramirez. 2012. GB. PMID:23270620
282	AGT1	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029527	ERR064926	Castillo-Ramirez. 2012. GB. PMID:23270620
283	AGT9	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029509	ERR064923	Castillo-Ramirez. 2012. GB. PMID:23270620
284	AGT67	1997	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029555	ERR064924	Castillo-Ramirez. 2012. GB. PMID:23270620
285	AGT120	1998	Argentina	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029554	ERR064919	Castillo-Ramirez. 2012. GB. PMID:23270620
286	ANS46	1982	Australia	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029537	ERR064898	Castillo-Ramirez. 2012. GB. PMID:23270620
287	BK2421	1996	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029510	ERR064899	Castillo-Ramirez. 2012. GB. PMID:23270620
288	BRA2	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029514	ERR064914	Castillo-Ramirez. 2012. GB. PMID:23270620

289	BRA36	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029559	ERR064917	Castillo-Ramirez. 2012. GB. PMID:23270620
290	BZ48	1997	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029508	ERR064918	Castillo-Ramirez. 2012. GB. PMID:23270620
291	CHI59	1998	China	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029524	ERR064937	Castillo-Ramirez. 2012. GB. PMID:23270620
292	CHI61	1998	China	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029538	ERR064938	Castillo-Ramirez. 2012. GB. PMID:23270620
293	CHL1	1997	Chile	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029526	ERR064915	Castillo-Ramirez. 2012. GB. PMID:23270620
294	CHL151	1998	Chile	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029539	ERR064916	Castillo-Ramirez. 2012. GB. PMID:23270620
295	D71	1996	Germany	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862562	ERR024980	Castillo-Ramirez. 2012. GB. PMID:23270620
296	D90	1996	Germany	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862585	ERR025801	Castillo-Ramirez. 2012. GB. PMID:23270620
297	DEN907	2001	Denmark	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029548	ERR064948	Castillo-Ramirez. 2012. GB. PMID:23270620
298	DEU2	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862577	ERR026652	Castillo-Ramirez. 2012. GB. PMID:23270620
299	DEU3	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862574	ERR026653	Castillo-Ramirez. 2012. GB. PMID:23270620
300	DEU5	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862580	ERR026655	Castillo-Ramirez. 2012. GB. PMID:23270620
301	DEU6	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862581	ERR026656	Castillo-Ramirez. 2012. GB. PMID:23270620
302	DEU8	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862579	ERR026658	Castillo-Ramirez. 2012. GB. PMID:23270620
303	DEU10	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862583	ERR026649	Castillo-Ramirez. 2012. GB. PMID:23270620
304	DEU11	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862554	ERR026650	Castillo-Ramirez. 2012. GB. PMID:23270620
305	DEU12	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862553	ERR026651	Castillo-Ramirez. 2012. GB. PMID:23270620
306	DEU14	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862555	ERR026664	Castillo-Ramirez. 2012. GB. PMID:23270620
307	DEU15	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862558	ERR026665	Castillo-Ramirez. 2012. GB. PMID:23270620
308	DEU16	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862557	ERR026666	Castillo-Ramirez. 2012. GB. PMID:23270620
309	DEU17	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862559	ERR026667	Castillo-Ramirez. 2012. GB. PMID:23270620
310	DEU19	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862561	ERR026668	Castillo-Ramirez. 2012. GB. PMID:23270620
311	DEU20	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862560	ERR026669	Castillo-Ramirez. 2012. GB. PMID:23270620
312	DEU23	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862541	ERR026662	Castillo-Ramirez. 2012. GB. PMID:23270620
313	DEU25	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862542	ERR026663	Castillo-Ramirez. 2012. GB. PMID:23270620
314	DEU27	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862536	ERR026677	Castillo-Ramirez. 2012. GB. PMID:23270620
315	DEU28	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862537	ERR026678	Castillo-Ramirez. 2012. GB. PMID:23270620
316	DEU29	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862538	ERR026679	Castillo-Ramirez. 2012. GB. PMID:23270620
317	DEU30	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862531	ERR026680	Castillo-Ramirez. 2012. GB. PMID:23270620
318	DEU35	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862662	ERR026683	Castillo-Ramirez. 2012. GB. PMID:23270620
319	DEU36	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862661	ERR026684	Castillo-Ramirez. 2012. GB. PMID:23270620
320	DEU37	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862660	ERR026674	Castillo-Ramirez. 2012. GB. PMID:23270620
321	DEU38	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862667	ERR026675	Castillo-Ramirez. 2012. GB. PMID:23270620
322	DEU39	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862666	ERR026676	Castillo-Ramirez. 2012. GB. PMID:23270620
323	DEU41	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862664	ERR026690	Castillo-Ramirez. 2012. GB. PMID:23270620
324	DEU42	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862655	ERR026691	Castillo-Ramirez. 2012. GB. PMID:23270620
325	DEU43	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862654	ERR026692	Castillo-Ramirez. 2012. GB. PMID:23270620
326	DEU46	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862570	ERR026694	Castillo-Ramirez. 2012. GB. PMID:23270620
327	DEU47	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862573	ERR026695	Castillo-Ramirez. 2012. GB. PMID:23270620
328	DEU49	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862567	ERR026697	Castillo-Ramirez. 2012. GB. PMID:23270620
329	DEU50	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862566	ERR026687	Castillo-Ramirez. 2012. GB. PMID:23270620
330	ES26	1996	Spain	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862530	ERR024981	Castillo-Ramirez. 2012. GB. PMID:23270620
331	FFP103	1990	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029533	ERR064932	Castillo-Ramirez. 2012. GB. PMID:23270620
332	GRE4	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029513	ERR064904	Castillo-Ramirez. 2012. GB. PMID:23270620
333	GRE18	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029546	ERR064902	Castillo-Ramirez. 2012. GB. PMID:23270620
334	GRE108	1998	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029521	ERR064936	Castillo-Ramirez. 2012. GB. PMID:23270620
335	GRE317	1999	Greece	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029564	ERR064903	Castillo-Ramirez. 2012. GB. PMID:23270620
336	H202	2006	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862584	ERR025802	Castillo-Ramirez. 2012. GB. PMID:23270620
337	H211	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862528	ERR024972	Castillo-Ramirez. 2012. GB. PMID:23270620
338	H216	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862587	ERR025803	Castillo-Ramirez. 2012. GB. PMID:23270620
339	H24	2005	Egypt	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862527	ERR024973	Castillo-Ramirez. 2012. GB. PMID:23270620

340	H482	2007	Romania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862595	ERR024983	Castillo-Ramirez. 2012. GB. PMID:23270620
341	HDG2	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029561	ERR064930	Castillo-Ramirez. 2012. GB. PMID:23270620
342	HGSA9	1997	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029534	ERR064921	Castillo-Ramirez. 2012. GB. PMID:23270620
343	HGSA142	2003	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029517	ERR064922	Castillo-Ramirez. 2012. GB. PMID:23270620
344	HS10	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029520	ERR064931	Castillo-Ramirez. 2012. GB. PMID:23270620
345	HS11	1992	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029542	ERR064929	Castillo-Ramirez. 2012. GB. PMID:23270620
346	HSJ216	1997	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029550	ERR064920	Castillo-Ramirez. 2012. GB. PMID:23270620
347	HU106	1996	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029518	ERR064927	Castillo-Ramirez. 2012. GB. PMID:23270620
348	HU109	1996	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029567	ERR064909	Castillo-Ramirez. 2012. GB. PMID:23270620
349	HU4	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862618	ERR026705	Castillo-Ramirez. 2012. GB. PMID:23270620
350	HU5	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862625	ERR026706	Castillo-Ramirez. 2012. GB. PMID:23270620
351	HU6	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862624	ERR026707	Castillo-Ramirez. 2012. GB. PMID:23270620
352	HU7	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862623	ERR026708	Castillo-Ramirez. 2012. GB. PMID:23270620
353	HU8	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862622	ERR026709	Castillo-Ramirez. 2012. GB. PMID:23270620
354	HU9	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862627	ERR026710	Castillo-Ramirez. 2012. GB. PMID:23270620
355	HU11	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862634	ERR026701	Castillo-Ramirez. 2012. GB. PMID:23270620
356	HU13	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862636	ERR024668	Castillo-Ramirez. 2012. GB. PMID:23270620
357	HU14	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862637	ERR024672	Castillo-Ramirez. 2012. GB. PMID:23270620
358	HU15	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862630	ERR024673	Castillo-Ramirez. 2012. GB. PMID:23270620
359	HU16	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862631	ERR024674	Castillo-Ramirez. 2012. GB. PMID:23270620
360	HU17	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862632	ERR024675	Castillo-Ramirez. 2012. GB. PMID:23270620
361	HU21	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862639	ERR024678	Castillo-Ramirez. 2012. GB. PMID:23270620
362	HU23	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862643	ERR024679	Castillo-Ramirez. 2012. GB. PMID:23270620
363	HU25	1993	Brazil	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029569	ERR024670	Castillo-Ramirez. 2012. GB. PMID:23270620
364	HU26	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862644	ERR024671	Castillo-Ramirez. 2012. GB. PMID:23270620
365	HUR18	1997	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029551	ERR064910	Castillo-Ramirez. 2012. GB. PMID:23270620
366	HUSA304	1993	Hungary	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029531	ERR064928	Castillo-Ramirez. 2012. GB. PMID:23270620
367	ICP5011	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029545	ERR064933	Castillo-Ramirez. 2012. GB. PMID:23270620
368	ICP5014	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029543	ERR064934	Castillo-Ramirez. 2012. GB. PMID:23270620
369	ICP5062	1993	Portugal	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029562	ERR064935	Castillo-Ramirez. 2012. GB. PMID:23270620
370	IU1	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862647	ERR024681	Castillo-Ramirez. 2012. GB. PMID:23270620
371	IU2	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862646	ERR024685	Castillo-Ramirez. 2012. GB. PMID:23270620
372	IU4	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862648	ERR024687	Castillo-Ramirez. 2012. GB. PMID:23270620
373	IU6	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862651	ERR024688	Castillo-Ramirez. 2012. GB. PMID:23270620
374	IU7	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862650	ERR024689	Castillo-Ramirez. 2012. GB. PMID:23270620
375	IU9	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862617	ERR024691	Castillo-Ramirez. 2012. GB. PMID:23270620
376	IU10	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862652	ERR024692	Castillo-Ramirez. 2012. GB. PMID:23270620
377	IU11	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862653	ERR024682	Castillo-Ramirez. 2012. GB. PMID:23270620
378	IU12	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862658	ERR024683	Castillo-Ramirez. 2012. GB. PMID:23270620
379	IU13	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862659	ERR024684	Castillo-Ramirez. 2012. GB. PMID:23270620
380	IU15	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862657	ERR024974	Castillo-Ramirez. 2012. GB. PMID:23270620
381	LHH1	1994	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029523	ERR064900	Castillo-Ramirez. 2012. GB. PMID:23270620
382	LIT68	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862586	ERR025804	Castillo-Ramirez. 2012. GB. PMID:23270620
383	LIT76	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862593	ERR025805	Castillo-Ramirez. 2012. GB. PMID:23270620
384	LIT89	1996	Lithuania	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862592	ERR025796	Castillo-Ramirez. 2012. GB. PMID:23270620
385	M116	2004	Vietnam	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862545	ERR024988	Castillo-Ramirez. 2012. GB. PMID:23270620
386	M278	2005	Portugal	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862548	ERR024991	Castillo-Ramirez. 2012. GB. PMID:23270620
387	M418	2006	India	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862549	ERR024992	Castillo-Ramirez. 2012. GB. PMID:23270620
388	M592	2006	Syria	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862550	ERR024993	Castillo-Ramirez. 2012. GB. PMID:23270620
389	M705	2007	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862613	ERR024958	Castillo-Ramirez. 2012. GB. PMID:23270620
390	M996	2008	China	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862551	ERR024994	Castillo-Ramirez. 2012. GB. PMID:23270620

391	M1229	2009	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862546	ERR024989	Castillo-Ramirez. 2012. GB. PMID:23270620
392	MAL1	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862552	ERR024984	Castillo-Ramirez. 2012. GB. PMID:23270620
393	MAL3	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862609	ERR024964	Castillo-Ramirez. 2012. GB. PMID:23270620
394	MAL9	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862607	ERR024966	Castillo-Ramirez. 2012. GB. PMID:23270620
395	MAL11	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862543	ERR024985	Castillo-Ramirez. 2012. GB. PMID:23270620
396	MAL119	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862611	ERR024962	Castillo-Ramirez. 2012. GB. PMID:23270620
397	MAL215	1996	Malaysia	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862608	ERR024963	Castillo-Ramirez. 2012. GB. PMID:23270620
398	MU1	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862569	ERR026688	Castillo-Ramirez. 2012. GB. PMID:23270620
399	MU3	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862565	ERR026003	Castillo-Ramirez. 2012. GB. PMID:23270620
400	MU4	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862564	ERR026007	Castillo-Ramirez. 2012. GB. PMID:23270620
401	MU5	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862604	ERR026008	Castillo-Ramirez. 2012. GB. PMID:23270620
402	MU6	2008	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862605	ERR026009	Castillo-Ramirez. 2012. GB. PMID:23270620
403	MU7	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862602	ERR026010	Castillo-Ramirez. 2012. GB. PMID:23270620
404	MU9	2007	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862603	ERR026011	Castillo-Ramirez. 2012. GB. PMID:23270620
405	MU10	2006	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862600	ERR026012	Castillo-Ramirez. 2012. GB. PMID:23270620
406	MU20	2009	Turkey	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862620	ERR026703	Castillo-Ramirez. 2012. GB. PMID:23270620
407	Na21	2006	Sri Lanka	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862544	ERR024986	Castillo-Ramirez. 2012. GB. PMID:23270620
408	NA32	2006	Denmark	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862615	ERR024967	Castillo-Ramirez. 2012. GB. PMID:23270620
409	P32	1996	Poland	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862589	ERR025795	Castillo-Ramirez. 2012. GB. PMID:23270620
410	R35	1987	United States	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029532	ERR064901	Castillo-Ramirez. 2012. GB. PMID:23270620
411	RA3	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862588	ERR025798	Castillo-Ramirez. 2012. GB. PMID:23270620
412	RA6	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862591	ERR025799	Castillo-Ramirez. 2012. GB. PMID:23270620
413	RA7	1996	Argentina	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862616	ERR024968	Castillo-Ramirez. 2012. GB. PMID:23270620
414	S2	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029563	ERR064953	Castillo-Ramirez. 2012. GB. PMID:23270620
415	S7	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029556	ERR064946	Castillo-Ramirez. 2012. GB. PMID:23270620
416	S21	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029536	ERR064958	Castillo-Ramirez. 2012. GB. PMID:23270620
417	S24	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029553	ERR064957	Castillo-Ramirez. 2012. GB. PMID:23270620
418	S25	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029535	ERR064951	Castillo-Ramirez. 2012. GB. PMID:23270620
419	S26	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029511	ERR064949	Castillo-Ramirez. 2012. GB. PMID:23270620
420	S38	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029568	ERR064947	Castillo-Ramirez. 2012. GB. PMID:23270620
421	S39	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029558	ERR064959	Castillo-Ramirez. 2012. GB. PMID:23270620
422	S40	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029528	ERR064944	Castillo-Ramirez. 2012. GB. PMID:23270620
423	S42	2006	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029515	ERR064955	Castillo-Ramirez. 2012. GB. PMID:23270620
424	S71	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029541	ERR064943	Castillo-Ramirez. 2012. GB. PMID:23270620
425	S78	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029525	ERR064954	Castillo-Ramirez. 2012. GB. PMID:23270620
426	S81	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029565	ERR064956	Castillo-Ramirez. 2012. GB. PMID:23270620
427	S85	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029529	ERR064939	Castillo-Ramirez. 2012. GB. PMID:23270620
428	S87	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029512	ERR064941	Castillo-Ramirez. 2012. GB. PMID:23270620
429	S93	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029560	ERR064942	Castillo-Ramirez. 2012. GB. PMID:23270620
430	S97	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029522	ERR064950	Castillo-Ramirez. 2012. GB. PMID:23270620
431	S102	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029519	ERR064945	Castillo-Ramirez. 2012. GB. PMID:23270620
432	S106	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029547	ERR064952	Castillo-Ramirez. 2012. GB. PMID:23270620
433	S130	2007	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029516	ERR064940	Castillo-Ramirez. 2012. GB. PMID:23270620
434	TUR1	1996	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029552	ERR064906	Castillo-Ramirez. 2012. GB. PMID:23270620
435	TUR9	1995	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029530	ERR064905	Castillo-Ramirez. 2012. GB. PMID:23270620
436	TUR27	1996	Turkey	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029557	ERR064908	Castillo-Ramirez. 2012. GB. PMID:23270620
437	UK102	1996	United Kingdom	-	WGS Reads	Illumina Genome Analyzer II	SAMEA862590	ERR025800	Castillo-Ramirez. 2012. GB. PMID:23270620
438	URU34	1997	Uruguay	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029544	ERR064925	Castillo-Ramirez. 2012. GB. PMID:23270620
439	URU110	1998	Uruguay	-	WGS Reads	Illumina HiSeq 2000	SAMEA1029540	ERR064911	Castillo-Ramirez. 2012. GB. PMID:23270620
440	SR2852	2005	United States	-	WGS Reads	Illumina HiSeq 2000	SAMN02649763	SRR1177896	Alam. GenomeBiolEvol. 2014. PMID:24787619
441	031_01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800273	ERR029457	Hsu. 2015. GenBiol. PMID:25903077

442	362_05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800274	ERR029458	Hsu. 2015. GenBiol. PMID:25903077
443	058_08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800253	ERR029464	Hsu. 2015. GenBiol. PMID:25903077
444	352_05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800280	ERR029467	Hsu. 2015. GenBiol. PMID:25903077
445	131_09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800276	ERR029469	Hsu. 2015. GenBiol. PMID:25903077
446	009_10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800287	ERR029470	Hsu. 2015. GenBiol. PMID:25903077
447	134_09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800277	ERR029472	Hsu. 2015. GenBiol. PMID:25903077
448	135_09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800278	ERR029473	Hsu. 2015. GenBiol. PMID:25903077
449	049_05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800292	ERR029476	Hsu. 2015. GenBiol. PMID:25903077
450	CGH1_08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800297	ERR030155	Hsu. 2015. GenBiol. PMID:25903077
451	CGH1_10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800177	ERR030157	Hsu. 2015. GenBiol. PMID:25903077
452	CGH2_08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800181	ERR030159	Hsu. 2015. GenBiol. PMID:25903077
453	CGH3_10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800213	ERR030170	Hsu. 2015. GenBiol. PMID:25903077
454	CGH6_06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800382	ERR030173	Hsu. 2015. GenBiol. PMID:25903077
455	CGH4_09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800215	ERR030176	Hsu. 2015. GenBiol. PMID:25903077
456	CGH4_10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800216	ERR030177	Hsu. 2015. GenBiol. PMID:25903077
457	CGH5_09	2009	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800383	ERR030182	Hsu. 2015. GenBiol. PMID:25903077
458	012_00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800207	ERR030186	Hsu. 2015. GenBiol. PMID:25903077
459	011_00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800205	ERR030191	Hsu. 2015. GenBiol. PMID:25903077
460	003_06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800206	ERR030192	Hsu. 2015. GenBiol. PMID:25903077
461	106_06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800200	ERR030194	Hsu. 2015. GenBiol. PMID:25903077
462	100_06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800201	ERR030195	Hsu. 2015. GenBiol. PMID:25903077
463	106_07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800360	ERR030200	Hsu. 2015. GenBiol. PMID:25903077
464	339_07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800357	ERR030201	Hsu. 2015. GenBiol. PMID:25903077
465	157_04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800333	ERR030202	Hsu. 2015. GenBiol. PMID:25903077
466	019_01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800332	ERR030203	Hsu. 2015. GenBiol. PMID:25903077
467	002_00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800331	ERR030204	Hsu. 2015. GenBiol. PMID:25903077
468	020_01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800329	ERR030205	Hsu. 2015. GenBiol. PMID:25903077
469	101_07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800327	ERR030209	Hsu. 2015. GenBiol. PMID:25903077
470	030_05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800258	ERR030213	Hsu. 2015. GenBiol. PMID:25903077
471	255_04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800353	ERR030216	Hsu. 2015. GenBiol. PMID:25903077
472	259_04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800354	ERR030217	Hsu. 2015. GenBiol. PMID:25903077
473	263_04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800352	ERR030219	Hsu. 2015. GenBiol. PMID:25903077
474	264_04	2004	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800261	ERR030220	Hsu. 2015. GenBiol. PMID:25903077
475	028_01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800260	ERR030221	Hsu. 2015. GenBiol. PMID:25903077
476	069_02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800271	ERR030224	Hsu. 2015. GenBiol. PMID:25903077
477	006_00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800285	ERR030228	Hsu. 2015. GenBiol. PMID:25903077
478	067_02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800265	ERR030232	Hsu. 2015. GenBiol. PMID:25903077
479	068_02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800272	ERR030233	Hsu. 2015. GenBiol. PMID:25903077
480	071_03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800269	ERR030235	Hsu. 2015. GenBiol. PMID:25903077
481	078_05	2005	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800264	ERR030238	Hsu. 2015. GenBiol. PMID:25903077
482	086_03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800263	ERR030239	Hsu. 2015. GenBiol. PMID:25903077
483	009_00	2000	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800336	ERR030240	Hsu. 2015. GenBiol. PMID:25903077
484	TTSH1_01	2001	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800251	ERR030242	Hsu. 2015. GenBiol. PMID:25903077
485	TTSH2_07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800229	ERR030249	Hsu. 2015. GenBiol. PMID:25903077
486	TTSH1_10	2010	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800219	ERR030252	Hsu. 2015. GenBiol. PMID:25903077
487	TTSH2_03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800226	ERR030255	Hsu. 2015. GenBiol. PMID:25903077
488	TTSH2_08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800230	ERR030259	Hsu. 2015. GenBiol. PMID:25903077
489	TTSH3_07	2007	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800196	ERR030260	Hsu. 2015. GenBiol. PMID:25903077
490	TTSH3_08	2008	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800232	ERR030261	Hsu. 2015. GenBiol. PMID:25903077
491	TTSH3_03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800234	ERR030266	Hsu. 2015. GenBiol. PMID:25903077
492	TTSH4_02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800293	ERR030276	Hsu. 2015. GenBiol. PMID:25903077

493	TTSH4_03	2003	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800372	ERR030277	Hsu. 2015. GenBiol. PMID:25903077
494	TTSH5_02	2002	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800345	ERR031666	Hsu. 2015. GenBiol. PMID:25903077
495	TTSH5_06	2006	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA800346	ERR031670	Hsu. 2015. GenBiol. PMID:25903077
496	WAU_8	1982	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021823	ERR053014	Hsu. 2015. GenBiol. PMID:25903077
497	WAU_9	1985	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021807	ERR053015	Hsu. 2015. GenBiol. PMID:25903077
498	WAU_10	1985	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021856	ERR053016	Hsu. 2015. GenBiol. PMID:25903077
499	WAU_16	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1022012	ERR053022	Hsu. 2015. GenBiol. PMID:25903077
500	WAU_17	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021973	ERR053023	Hsu. 2015. GenBiol. PMID:25903077
501	WAU_18	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021952	ERR053024	Hsu. 2015. GenBiol. PMID:25903077
502	WAU_19	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021934	ERR053025	Hsu. 2015. GenBiol. PMID:25903077
503	WAU_21	1996	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021910	ERR053027	Hsu. 2015. GenBiol. PMID:25903077
504	WAU_26	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021849	ERR053032	Hsu. 2015. GenBiol. PMID:25903077
505	WAU_29	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021995	ERR053035	Hsu. 2015. GenBiol. PMID:25903077
506	WAU_31	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021970	ERR053037	Hsu. 2015. GenBiol. PMID:25903077
507	WAU_34	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021933	ERR053040	Hsu. 2015. GenBiol. PMID:25903077
508	WAU_39	1997	Singapore	-	WGS Reads	Illumina Genome Analyzer II	SAMEA1021951	ERR053044	Hsu. 2015. GenBiol. PMID:25903077
509	T9_N3	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958822	ERR023888	Tong. GenRes. 2015. PMID:25491771
510	T234_W2	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958806	ERR023872	Tong. GenRes. 2015. PMID:25491771
511	T71_N2	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958988	ERR033836	Tong. GenRes. 2015. PMID:25491771
512	T99_N5	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958963	ERR033828	Tong. GenRes. 2015. PMID:25491771
513	T12_N1	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958848	ERR023884	Tong. GenRes. 2015. PMID:25491771
514	T271_C1	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958810	ERR023867	Tong. GenRes. 2015. PMID:25491771
515	T35_N2	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958775	ERR033823	Tong. GenRes. 2015. PMID:25491771
516	T95_N2	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958962	ERR033833	Tong. GenRes. 2015. PMID:25491771
517	T192_T3	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958900	ERR023907	Tong. GenRes. 2015. PMID:25491771
518	T303_N3	2008	Thailand	-	WGS Reads	Illumina Genome Analyzer II	SAMEA958836	ERR033080	Tong. GenRes. 2015. PMID:25491771
519	SA13002	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03856026	SRR2097047	Kong. 2016. PLoS One. PMID:26950298
520	SA13005	2013	China	-	WGS Reads	Illumina MiSeq	SAMN03878399	SRR2106469	Kong. 2016. PLoS One. PMID:26950298
521	SA13007	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878400	SRR2106470	Kong. 2016. PLoS One. PMID:26950298
522	SA13009	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878401	SRR2106471	Kong. 2016. PLoS One. PMID:26950298
523	SA13012	2012	China	-	WGS Reads	Illumina MiSeq	SAMN03878402	SRR2106472	Kong. 2016. PLoS One. PMID:26950298
524	SA13023	2013	China	-	WGS Reads	Illumina MiSeq	SAMN03878403	SRR2106473	Kong. 2016. PLoS One. PMID:26950298
525	BSAC_st3021	2010	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464569	ERR124493	Reuter. GenRes. 2016. PMID:26672018
526	BSAC_st3046	2010	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464296	ERR124495	Reuter. GenRes. 2016. PMID:26672018
527	BSAC_st697	2003	England	-	WGS Reads	Illumina HiSeq 2000	SAMEA1464583	ERR129279	Reuter. GenRes. 2016. PMID:26672018
528	D8	2007	Gambia	-	WGS Reads	Illumina MiSeq	SAMEA3727072	ERR1213802	Senghore. AppEnvMicro. 2016. PMID:27474712
529	DS_009	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448838	ERR1069915	Moradigaravand. mBio. 2017. PMID: 28679748
530	DS_014	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448843	ERR1069920	Moradigaravand. mBio. 2017. PMID: 28679748
531	US_002	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448943	ERR1070030	Moradigaravand. mBio. 2017. PMID: 28679748
532	US_008	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448948	ERR1070035	Moradigaravand. mBio. 2017. PMID: 28679748
533	US_030	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448965	ERR1070052	Moradigaravand. mBio. 2017. PMID: 28679748
534	US_017	2011	Thailand	-	WGS Reads	Illumina HiSeq 2000	SAMEA3448955	ERR1070042	Moradigaravand. mBio. 2017. PMID: 28679748

Bibliography

1. Becker K, Skov RL, von Eiff C. 2015. *Staphylococcus*, *Micrococcus*, and Other Catalase-Positive Cocci, p 354. In Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, Warnock DW (ed), Manual of Clinical Microbiology, 11 ed, vol 1. ASM Press, Washington DC.
2. Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures G. 2017. Prokaryotic Nomenclature Up-to-Date. <http://www.dsmz.de/bacterial-diversity/prokaryotic-nomenclature-up-to-date>. Accessed
3. Stackebrandt E, Rainey FA, Ward-Rainey NL. 1997. Proposal for a new hierarchic classification system, *Actinobacteria* *Iclassis nov.* Int J Syst Bacteriol 47:479-491.
4. Garrity GM, Bell JA, Lilburn TG. 2004. Taxonomic outline of the procaryotes. In Garrity GM, Brenner DJ, Krieg NR, Staley JT (ed), Bergey's Manual of Systematic Bacteriology, 2 ed, vol 3. Springer-Verlag, New York, NY.
5. Ludwig W, Schleifer KH, Whitman WB. 2009. Revised road map to the phylum *Firmicutes*, p 1-13. In De Vos P, Garrity GM, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman WB (ed), Bergey's Manual of Systematic Bacteriology, 2 ed, vol 3. Springer, New York, NY.
6. Becker RE, Bubeck-Wardenburg J. 2015. *Staphylococcus aureus* and the skin: a longstanding and complex interaction. Skinmed 13:111-9; quiz 120.
7. Lowy FD. 1998. *Staphylococcus aureus* infections. N Engl J Med 339:520-532.
8. Peacock SJ, de Silva I, Lowy FD. 2001. What determines nasal carriage of *Staphylococcus aureus*? Trends Microbiol 9:605-10.
9. Nulens E, Gould I, MacKenzie F, Deplano A, Cookson B, Alp E, Bouza E, Voss A. 2005. *Staphylococcus aureus* carriage among participants at the 13th European Congress of Clinical Microbiology and Infectious Diseases. Eur J Clin Microbiol Infect Dis 24:145-8.
10. Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, Nouwen JL. 2005. The role of nasal carriage in *Staphylococcus aureus* infections. Lancet Infect Dis 5:751-62.
11. Sollid JU, Furberg AS, Hanssen AM, Johannessen M. 2014. *Staphylococcus aureus*: determinants of human carriage. Infect Genet Evol 21:531-41.
12. Ogston A. 1882. *Micrococcus* Poisoning. J Anat Physiol 16:526-67.
13. Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. 2015. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev 28:603-61.
14. Kaasch AJ, Barlow G, Edgeworth JD, Fowler VG, Jr., Hellmich M, Hopkins S, Kern WV, Llewelyn MJ, Rieg S, Rodriguez-Bano J, Scarborough M, Seifert H, Soriano A, Tilley R, Torok ME, Weiss V, Wilson AP, Thwaites GE, Isac ISU, Colleagues. 2014. *Staphylococcus aureus* bloodstream infection: a pooled analysis of five prospective, observational studies. J Infect 68:242-51.
15. Gordon RJ, Lowy FD. 2008. Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. Clin Infect Dis 46 Suppl 5:S350-9.
16. Kluytmans J, van Belkum A, Verbrugh H. 1997. Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev 10:505-20.
17. Skinner D, Keefer, C. S. 1941. Significance of bacteremia caused by *Staphylococcus aureus*. Arch Intern Med 68:851-875.

18. Wilson R, Hamburger M. 1957. Fifteen years' experience with *staphylococcus* septicemia in a large city hospital; analysis of fifty-five cases in the Cincinnati General Hospital 1940 to 1954. *Am J Med* 22:437-57.
19. Chambers HF, Deleo FR. 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* 7:629-41.
20. Peacock SJ, Paterson GK. 2015. Mechanisms of Methicillin Resistance in *Staphylococcus aureus*. *Annu Rev Biochem* 84:577-601.
21. Tong SY, van Hal SJ, Einsiedel L, Currie BJ, Turnidge JD, Australian New Zealand Cooperative on Outcomes in Staphylococcal S. 2012. Impact of ethnicity and socio-economic status on *Staphylococcus aureus* bacteremia incidence and mortality: a heavy burden in Indigenous Australians. *BMC Infect Dis* 12:249.
22. Williamson DA, Lim A, Thomas MG, Baker MG, Roberts SA, Fraser JD, Ritchie SR. 2013. Incidence, trends and demographics of *Staphylococcus aureus* infections in Auckland, New Zealand, 2001-2011. *BMC Infect Dis* 13:569.
23. Johnson AP, Pearson A, Duckworth G. 2005. Surveillance and epidemiology of MRSA bacteraemia in the UK. *J Antimicrob Chemother* 56:455-62.
24. Laupland KB, Ross T, Gregson DB. 2008. *Staphylococcus aureus* bloodstream infections: risk factors, outcomes, and the influence of methicillin resistance in Calgary, Canada, 2000-2006. *J Infect Dis* 198:336-43.
25. Mostofsky E, Lipsitch M, Regev-Yochay G. 2011. Is methicillin-resistant *Staphylococcus aureus* replacing methicillin-susceptible *S. aureus*? *J Antimicrob Chemother* 66:2199-214.
26. Holland TL, Arnold C, Fowler VG, Jr. 2014. Clinical management of *Staphylococcus aureus* bacteremia: a review. *JAMA* 312:1330-41.
27. van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB. 2012. Predictors of mortality in *Staphylococcus aureus* Bacteremia. *Clin Microbiol Rev* 25:362-86.
28. Collignon P, Nimmo GR, Gottlieb T, Gosbell IB, Australian Group on Antimicrobial R. 2005. *Staphylococcus aureus* bacteremia, Australia. *Emerg Infect Dis* 11:554-61.
29. Hill PC, Birch M, Chambers S, Drinkovic D, Ellis-Pegler RB, Everts R, Murdoch D, Pottumarthy S, Roberts SA, Swager C, Taylor SL, Thomas MG, Wong CG, Morris AJ. 2001. Prospective study of 424 cases of *Staphylococcus aureus* bacteraemia: determination of factors affecting incidence and mortality. *Intern Med J* 31:97-103.
30. Hill PC, Wong CG, Voss LM, Taylor SL, Pottumarthy S, Drinkovic D, Morris AJ. 2001. Prospective study of 125 cases of *Staphylococcus aureus* bacteremia in children in New Zealand. *Pediatr Infect Dis J* 20:868-73.
31. Tong SY, Holden MT, Nickerson EK, Cooper BS, Koser CU, Cori A, Jombart T, Cauchemez S, Fraser C, Wuthiekanun V, Thaipadungpanit J, Hongsuwan M, Day NP, Limmathurotsakul D, Parkhill J, Peacock SJ. 2015. Genome sequencing defines phylogeny and spread of methicillin-resistant *Staphylococcus aureus* in a high transmission setting. *Genome Res* 25:111-8.
32. McCaig LF, McDonald LC, Mandal S, Jernigan DB. 2006. *Staphylococcus aureus*-associated skin and soft tissue infections in ambulatory care. *Emerg Infect Dis* 12:1715-23.

33. Esposito S, Bassetti M, Concia E, De Simone G, De Rosa FG, Grossi P, Novelli A, Menichetti F, Petrosillo N, Tinelli M, Tumbarello M, Sanguinetti M, Viale P, Venditti M, Viscoli C, Italian Society of I, Tropical D. 2017. Diagnosis and management of skin and soft-tissue infections (SSTI). A literature review and consensus statement: an update. *J Chemother* 29:197-214.
34. Hayward A, Knott F, Petersen I, Livermore DM, Duckworth G, Islam A, Johnson AM. 2008. Increasing hospitalizations and general practice prescriptions for community-onset staphylococcal disease, England. *Emerg Infect Dis* 14:720-6.
35. Lautz TB, Raval MV, Barsness KA. 2011. Increasing national burden of hospitalizations for skin and soft tissue infections in children. *J Pediatr Surg* 46:1935-41.
36. Vaska VL, Nimmo GR, Jones M, Grimwood K, Paterson DL. 2012. Increases in Australian cutaneous abscess hospitalisations: 1999-2008. *Eur J Clin Microbiol Infect Dis* 31:93-6.
37. Williamson DA, Zhang J, Ritchie SR, Roberts SA, Fraser JD, Baker MG. 2014. *Staphylococcus aureus* infections in New Zealand, 2000-2011. *Emerg Infect Dis* 20:1156-61.
38. Williamson DA, Heffernan H. 2014. The changing landscape of antimicrobial resistance in New Zealand. *N Z Med J* 127:41-54.
39. Williamson DA, Monecke S, Heffernan H, Ritchie SR, Roberts SA, Upton A, Thomas MG, Fraser JD. 2014. High usage of topical fusidic acid and rapid clonal expansion of fusidic acid-resistant *Staphylococcus aureus*: a cautionary tale. *Clin Infect Dis* 59:1451-4.
40. Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, Cartwright RA, Simpson AJ, Rambaut A, Nubel U, Fitzgerald JR. 2009. Recent human-to-poultry host jump, adaptation and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci USA* 106:19545-50.
41. Fitzgerald JR, Holden MT. 2016. Genomics of Natural Populations of *Staphylococcus aureus*. *Annu Rev Microbiol* 70:459-78.
42. Vanderhaeghen W, Hermans K, Haesebrouck F, Butaye P. 2010. Methicillin-resistant *Staphylococcus aureus* (MRSA) in food production animals. *Epidemiol Infect* 138:606-25.
43. van Belkum A, Melles DC, Peeters JK, van Leeuwen WB, van Duinbergen E, Huijsdens XW, Spalburg E, de Neeling AJ, Verbrugh HA, Dutch Working Party on S, Research of M-S. 2008. Methicillin-resistant and -susceptible *Staphylococcus aureus* sequence type 398 in pigs and humans. *Emerg Infect Dis* 14:479-83.
44. Catry B, Van Duinbergen E, Pomba MC, Greko C, Moreno MA, Pyorala S, Ruzauskas M, Sanders P, Threlfall EJ, Ungemach F, Torneke K, Munoz-Madero C, Torren-Edo J, Scientific Advisory Group on A. 2010. Reflection paper on MRSA in food-producing and companion animals: epidemiology and control options for human and animal health. *Epidemiol Infect* 138:626-44.
45. Kuroda M, Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oquchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Marakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H., Hiramatsu, K. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* 357:1225-1240.

46. Z - Baba T, Takeuchi F, Kuroda M, Yazawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 359:1819-1827.
47. Holden MTG, Feil EJ, Lindsay JA, Peacock SJ, Day NPJ, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth CC, Tracey, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabbinoiwitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci USA* 101:9786-9791.
48. Gillaspay AF, Worrell V, Orvis J, Roe BA, Dyer DW, Iandolo JJ. 2006. *Staphylococcus aureus* NCTC 8325 genome, 2 ed. ASM Press, Washington, DC.
49. Gill SR, Fouts, D. E., Archer, G. L., Mongodin, E. F., Deboy, R. T., Ravel, J., Paulsen, I. T., Kolonay, J. F., Brinkac, L., Beanan, M., Dodson, R. J., Daugherty, S. C., Madupu, R., Angiuoli, S. V., Durkin, A. S., Haft, D. H., Vamathevan, J., Khouri, H., Utterback, T., Lee, C., Dimitrov, G., Jiang, L., Din, H., Weidman, J., Tran, K., Kang, K., Hance, I. R., Nelson, K. E., Fraser, C. M. 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J Bacteriol* 187:2426-2438.
50. Lindsay JA, Holden, M. T. G. 2004. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol* 12:378-385.
51. Chua K, Stinear TP, Howden BP. 2013. Functional genomics of *Staphylococcus aureus*. *Brief Funct Genomics*.
52. Lindsay JA, Holden, M. T. G. 2006. Understanding the rise of the superbug: investigation of the evolution and genomic variation of *Staphylococcus aureus*. *Funct Integr Genomics* 6:186-201.
53. Malachowa N, DeLeo, F. R. 2010. Mobile genetic elements of *Staphylococcus aureus*. *Cell Mol Life Sci* 67:3057-3071.
54. Partridge SR, Kwong SM, Firth N, Jensen SO. 2018. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* 31.
55. McCarthy AJ, Lindsay JA. 2012. The distribution of plasmids that carry virulence and resistance genes in *Staphylococcus aureus* is lineage associated. *BMC Microbiol* 12:104.
56. Petit RA, 3rd, Read TD. 2018. *Staphylococcus aureus* viewed from the perspective of 40,000+ genomes. *PeerJ* 6:e5261.
57. Musser JM, Kapur, V. 1992. Clonal analysis of methicillin-resistant *Staphylococcus aureus* strain from intercontinental sources: association of the *mec* gene with divergent phylogenetic lineages implies dissemination by horizontal transfer and recombination. *J Clin Micro* 30:2058-2063.
58. Kreiswirth B, Kornblum, J., Arbeit, R. D., Eisner, W., Maslow, J. N., McGeer, A. Low, D. E., Novick, R. P. 1993. Evidence for a clonal origin of methicillin resistance in *Staphylococcus aureus*. *Science* 259:227-230.

59. Murray PR, Rosenthal, K. S., Pfaller, M. A. 2009. Medical Microbiology, 6 ed. Mosby Elsevier, Philadelphia, USA.
60. Fitzgerald JR, Sturdevant, D. E., Mackie, S. M., Gill, S. R. & Musser, J. M. 2001. Evolutionary genomics of *Staphylococcus aureus*: Insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. PNAS 98:8821-8826.
61. Strommenger B, Bräulke C, Heuck D, Schmidt C, Pasemann B, Nübel U, Witte W. 2008. spa Typing of *Staphylococcus aureus* as a frontline tool in epidemiological typing. J Clin Microbiol 46:574-81.
62. Goering RV, Morrison D, Al-Doori Z, Edwards GF, Gemmell CG. 2008. Usefulness of *mec*-associated direct repeat unit (dru) typing in the epidemiological analysis of highly clonal methicillin-resistant *Staphylococcus aureus* in Scotland. Clin Microbiol Infect 14:964-9.
63. Ryffel C, Bucher R, Kayser FH, Berger-Bachi B. 1991. The *Staphylococcus aureus mec* determinant comprises an unusual cluster of direct repeats and codes for a gene product similar to the Escherichia coli sn-glycerophosphoryl diester phosphodiesterase. J Bacteriol 173:7416-22.
64. Herschleb J, Ananiev G, Schwartz DC. 2007. Pulsed-field gel electrophoresis. Nat Protoc 2:677-84.
65. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 95:3140-5.
66. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol 38:1008-15.
67. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci USA 99:7687-7692.
68. Feil EJ, Li, B. C., Aanensen, D. M., Hanage, W. P., Spratt, B. G. 2004. eBURST: Inferring patterns of evolutionary descent among clustres of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186:1518-1530.
69. Aanensen DM, Feil EJ, Holden MT, Dordel J, Yeats CA, Fedosejev A, Goater R, Castillo-Ramirez S, Corander J, Colijn C, Chlebowicz MA, Schouls L, Heck M, Pluister G, Ruimy R, Kahlmeter G, Ahman J, Matuschek E, Friedrich AW, Parkhill J, Bentley SD, Spratt BG, Grundmann H, European SRLWG. 2016. Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive *Staphylococcus aureus* in Europe. MBio 7.
70. Young BC, Earle SG, Soeng S, Sar P, Kumar V, Hor S, Sar V, Bousfield R, Sanderson ND, Barker L, Stoesser N, Emary KR, Parry CM, Nickerson EK, Turner P, Bowden R, Crook DW, Wyllie DH, Day NP, Wilson DJ, Moore CE. 2019. Panton-Valentine leucocidin is the key determinant of *Staphylococcus aureus* pyomyositis in a bacterial GWAS. Elife 8.
71. Lyon BR, Skurray, R. 1987. Antimicrobial resistance of *Staphylococcus aureus*: genetic basis. Microbiol Review 51:88-134.
72. Clatworthy AE, Pierson, E., Hung, D. T. 2007. Targeting virulence: a new paradigm for antimicrobial therapy. Nat Chem Biol 3:541-548.

73. McGuinness WA, Malachowa N, DeLeo FR. 2017. Vancomycin Resistance in *Staphylococcus aureus*. *Yale J Biol Med* 90:269-281.
74. Foster TJ. 2017. Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. *FEMS Microbiol Rev* 41:430-449.
75. Harkins CP, Pichon B, Doumith M, Parkhill J, Westh H, Tomasz A, de Lencastre H, Bentley SD, Kearns AM, Holden MTG. 2017. Methicillin-resistant *Staphylococcus aureus* emerged long before the introduction of methicillin into clinical practice. *Genome Biol* 18:130.
76. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schaberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K. 2015. A new antibiotic kills pathogens without detectable resistance. *Nature* 517:455-9.
77. Yousefi M, Fallah F, Arshadi M, Pourmand MR, Hashemi A, Pourmand G. 2017. Identification of tigecycline- and vancomycin-resistant *Staphylococcus aureus* strains among patients with urinary tract infection in Iran. *New Microbes New Infect* 19:8-12.
78. Kumar M. 2016. Multidrug-Resistant *Staphylococcus aureus*, India, 2013-2015. *Emerg Infect Dis* 22:1666-7.
79. Chopra I, Roberts M. 2001. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* 65:232-60 ; second page, table of contents.
80. Harris DM. 1973. Staphylococcal infection in an intensive-care unit, and its relation to infection in the remainder of the hospital. *J Hyg (Lond)* 71:341-8.
81. Evans RJ, Waterworth PM. 1966. Naturally-occurring fusidic acid resistance in staphylococci and its linkage to other resistances. *J Clin Pathol* 19:555-60.
82. Maccabe AF, Gould JC. 1956. The epidemiology of an erythromycin resistant *Staphylococcus*. *Scott Med J* 1:223-6.
83. Lacey RW, Mitchell AA. 1969. Gentamicin-resistant *Staphylococcus aureus*. *Lancet* 2:1425-6.
84. Mangili A, Bica I, Snyderman DR, Hamer DH. 2005. Daptomycin-resistant, methicillin-resistant *Staphylococcus aureus* bacteremia. *Clin Infect Dis* 40:1058-60.
85. Nakhla LS. 1972. Resistance of *Staphylococcus aureus* to sulphamethoxazole and trimethoprim. *J Clin Pathol* 25:708-12.
86. Emmerson AM, Jones AM. 2003. The quinolones: decades of development and use. *J Antimicrob Chemother* 51 Suppl 1:13-20.
87. Humphreys H, Mulvihill E. 1985. Ciprofloxacin-resistant *Staphylococcus aureus*. *Lancet* 2:383.
88. US Department of Health and Human Services Centers for Disease Control and Prevention. 2013. Antibiotic Resistance Threats in the United States. <https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf>. Accessed
89. Munita JM, Arias CA. 2016. Mechanisms of Antibiotic Resistance. *Microbiol Spectr* 4.
90. Guerillot R, Li L, Baines S, Howden B, Schultz MB, Seemann T, Monk I, Pidot SJ, Gao W, Giulieri S, Goncalves da Silva A, D'Agata A, Tomita T, Peleg AY, Stinear TP, Howden BP. 2018. Comprehensive antibiotic-linked mutation assessment by resistance mutation sequencing (RM-seq). *Genome Med* 10:63.

91. Lowy FD. 2003. Antimicrobial resistance: the example of *Staphylococcus aureus*. J Clin Invest 111:1265-73.
92. Howden BP, Stinear TP, Davies JK. 2008. Complete genome sequencing of *Staphylococcus aureus*: insights into virulence and antimicrobial resistance. Microbio Australia 29:115-20.
93. Robinson DA, Enright MC. 2004. Evolution of *Staphylococcus aureus* by large chromosomal replacements. J Bacteriol 186:1060-1064.
94. Dengler V, Meier PS, Heusser R, Berger-Bachi B, McCallum N. 2011. Induction kinetics of the *Staphylococcus aureus* cell wall stress stimulon in response to different cell wall active antibiotics. BMC Microbiol 11:16.
95. McCallum N, Berger-Bachi B, Senn MM. 2010. Regulation of antibiotic resistance in *Staphylococcus aureus*. Int J Med Microbiol 300:118-29.
96. Cho H, Uehara T, Bernhardt TG. 2014. Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. Cell 159:1300-11.
97. Giesbrecht P, Kersten T, Maidhof H, Wecke J. 1998. Staphylococcal cell wall: morphogenesis and fatal variations in the presence of penicillin. Microbiol Mol Biol Rev 62:1371-414.
98. Ghuysen JM. 1994. Molecular structures of penicillin-binding proteins and beta-lactamases. Trends Microbiol 2:372-80.
99. Rammelkamp CH, Maxon T. 1942. Resistance of *Staphylococcus aureus* to the Action of Penicillin. Proc Soc Exp Biol and Med 51:386-389.
100. Bondi A, Jr., Dietz CC. 1945. Penicillin resistant staphylococci. Proc Soc Exp Biol Med 60:55-8.
101. Kirby WMM. 1944. Extraction of a highly potent penicillin inactivator from penicillin resistant *Staphylococci*. Science 99:452-453.
102. Zhang HZ, Hackbarth, C. J., Chansky, K. M., Chambers, H. F. 2001. A proteolytic transmembrane signaling pathway and resistance to β -lactams in *Staphylococci*. Science 291:1962-1965.
103. Zygmunt DJ, Stratton CW, Kernodle DS. 1992. Characterization of four beta-lactamases produced by *Staphylococcus aureus*. Antimicrob Agents Chemother 36:440-5.
104. Jensen SO, Lyon BR. 2009. Genetics of antimicrobial resistance in *Staphylococcus aureus*. Future Microbiol 4:565-82.
105. Crane JK. 2014. Resurgence of penicillin-susceptible *Staphylococcus aureus* at a hospital in New York State, USA. J Antimicrob Chemother 69:280-1.
106. Richmond MH. 1979. Beta-lactam antibiotics and beta-lactamases: two sides of a continuing story. Rev Infect Dis 1:30-8.
107. Jevons MP. 1961. Celbenin-resistant *Staphylococci*. Br Med J 1:124-125.
108. Jevons MP, Coe AW, Parker MT. 1963. Methicillin resistance in *Staphylococci*. Lancet:904-907.
109. Barber M. 1961. Methicillin-resistant *staphylococci*. J Clin Pathol 14:385-93.
110. Rolinson GN, Stevens S, Batchelor FR, Wood JC, Chain EB. 1960. Bacteriological studies on a new penicillin-BRL. 1241. Lancet 2:564-7.
111. Hartman BJ, Tomasz A. 1984. Low-affinity penicillin-binding protein associated with beta-lactam resistance in *Staphylococcus aureus*. J Bacteriol 158:513-6.

112. Utsui Y, Yokota T. 1985. Role of an altered penicillin-binding protein in methicillin- and cephem-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 28:397-403.
113. Lim D, Strynadka, C. J. 2002. Structural basis for the B-lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus*. *Nat Struct Biol* 9:870-876.
114. Pinho MG, Filipe SR, de Lencastre H, Tomasz A. 2001. Complementation of the essential peptidoglycan transpeptidase function of penicillin-binding protein 2 (PBP2) by the drug resistance protein PBP2A in *Staphylococcus aureus*. *J Bacteriol* 183:6525-31.
115. Ito T, Katayama, Y., Asada, K., Mori, N., Tsutsumimoto, K., Tiensasitorn, C., Hiramatsu, K. 2001. Structural comparison of three types of Staphylococcal Cassette Chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 45:1323-1336.
116. Ito T, Katayama, Y., Hiramatsu, K. 1999. Cloning and nucleotide sequencing determination of the entire *mec* DNA of pre-methicillin-resistant *Staphylococcus aureus* N315. *Antimicrob Agents Chemother* 43:1449-1458.
117. Hiramatsu K, Katayama Y, Matsuo M, Sasaki T, Morimoto Y, Sekiguchi A, Baba T. 2014. Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *J Infect Chemother* 20:593-601.
118. Andreis SN, Perreten V, Schwendener S. 2017. Novel beta-Lactamase blaARL in *Staphylococcus arlettae*. *mSphere* 2.
119. Paterson GK, Harrison EM, Holmes MA. 2014. The emergence of *mecC* methicillin-resistant *Staphylococcus aureus*. *Trends Microbiol* 22:42-7.
120. Tsubakishita S, Kuwahara-Arai K, Baba T, Hiramatsu K. 2010. Staphylococcal cassette chromosome *mec*-like element in *Macroccoccus caseolyticus*. *Antimicrob Agents Chemother* 54:1469-75.
121. Becker K, van Alen S, Idelevich EA, Schleimer N, Seggewiss J, Mellmann A, Kaspar U, Peters G. 2018. Plasmid-Encoded Transferable *mecB*-Mediated Methicillin Resistance in *Staphylococcus aureus*. *Emerg Infect Dis* 24:242-248.
122. Maslanova I, Wertheimer Z, Sedlacek I, Svec P, Indrakova A, Kovarovic V, Schumann P, Sproer C, Kralova S, Sedo O, Kristofova L, Vrbovska V, Fuzik T, Petras P, Zdrahal Z, Ruzickova V, Doskar J, Pantucek R. 2018. Description and Comparative Genomics of *Macroccoccus caseolyticus* subsp. *hominis* subsp. nov., *Macroccoccus goetzii* sp. nov., *Macroccoccus epidermidis* sp. nov., and *Macroccoccus bohemicus* sp. nov., Novel Macrococci From Human Clinical Material With Virulence Potential and Suspected Uptake of Foreign DNA by Natural Transformation. *Front Microbiol* 9:1178.
123. Parker MT, Jevons MP. 1964. A survey of methicillin resistance in *Staphylococcus aureus*. *Postgrad Medical Journal* 40:170-178.
124. Benner EJ, Kayser FH. 1968. Growing clinical significance of methicillin-resistant *Staphylococcus aureus*. *Lancet* 292:741-744.
125. Parker MT, Hewitt JH. 1970. Methicillin resistance in *Staphylococcus aureus*. *Lancet*:800-804.
126. Rosendal K, Bulow, P., Bentzon, M. W., Eriksen, K. R. 1976. *Staphylococcus aureus* strains isolated in danish hospitals from January 1st, 1966 to December 31st, 1974. *Acta Path Microbiol Scand* 84B:359-368.

127. Crisostomo MI, Westh H, Tomasz A, Chung M, Oliveira DC, de Lencastre H. 2001. The evolution of methicillin resistance in *Staphylococcus aureus*: similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. *Proc Natl Acad Sci U S A* 98:9865-70.
128. Barrett FF, McGehee RF, Jr., Finland M. 1968. Methicillin-resistant *Staphylococcus aureus* at Boston City Hospital. Bacteriologic and epidemiologic observations. *N Engl J Med* 279:441-8.
129. Bran JL, Levison ME, Kaye D. 1972. Survey for methicillin-resistant staphylococci. *Antimicrob Agents Chemother* 1:235-6.
130. Rountree PM, Beard MA. 1968. Hospital strain of *Staphylococcus aureus* with particular reference to methicillin-resistant strains. *Med J Aust* 2:1163-1168.
131. Udo EE, Pearman, J. W., Grubb, W. B. 1993. Genetic analysis of community isolates of methicillin-resistant *Staphylococcus aureus* in Western Australia. *J Hosp Infect* 25:97-108.
132. Chua K, Laurent F, Coombs G, Grayson ML, Howden BP. 2011. Not community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA)! A clinician's guide to community MRSA - Its evolving antimicrobial resistance and implication for therapy. *Clin Infect Dis* 52:99-114.
133. Francisco AP, Bugalho, M., Ramirez, M., Carrico, J. A. 2009. Global optimal eBURST analysis of multilocus typing data using a graphic matriod approach. *BMC Bioinformatics* 10:152.
134. Monecke S, Jatzwauk L, Muller E, Nitschke H, Pfohl K, Slickers P, Reissig A, Ruppelt-Lorz A, Ehricht R. 2016. Diversity of SCCmec Elements in *Staphylococcus aureus* as Observed in South-Eastern Germany. *PLoS One* 11:e0162654.
135. Hanssen AM, Ericson Sollid JU. 2006. SCCmec in staphylococci: genes on the move. *FEMS Immunol Med Microbiol* 46:8-20.
136. Rudkin JK, Edwards AM, Bowden MG, Brown EL, Pozzi C, Waters EM, Chan WC, Williams P, O'Gara JP, Massey RC. 2012. Methicillin resistance reduces the virulence of healthcare-associated methicillin-resistant *Staphylococcus aureus* by interfering with the *agr* quorum sensing system. *J Infect Dis* 205:798-806.
137. Baines SL, Holt KE, Schultz MB, Seemann T, Howden BO, Jensen SO, van Hal SJ, Coombs GW, Firth N, Powell DR, Stinear TP, Howden BP. 2015. Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*. *MBio* 6:e00080.
138. Centers for Disease Control and Prevention. 1999. From the Centers for Disease Control and Prevention. Four pediatric deaths from community-acquired methicillin-resistant *Staphylococcus aureus*--Minnesota and North Dakota, 1997-1999. *JAMA* 282:1123-5.
139. Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 359:1819-27.
140. Stegger M, Wirth T, Andersen PS, Skov RL, De Grassi A, Simoes PM, Tristan A, Petersen A, Aziz M, Kiil K, Cirkovic I, Udo EE, del Campo R, Vuopio-Varkila J, Ahmad N, Tokajian S, Peters G, Schaumburg F, Olsson-Liljequist B, Givskov M, Driebe EE, Vigh HE, Shittu A, Ramdani-Bougessa N, Rasigade JP, Price LB, Vandenesch F, Larsen AR, Laurent F. 2014. Origin and evolution of European community-acquired methicillin-resistant *Staphylococcus aureus*. *MBio* 5:e01044-14.

141. DeLeo FR, Otto M, Kreiswirth BN, Chambers HF. 2010. Community-associated methicillin-resistant *Staphylococcus aureus*. *Lancet* 375:1557-68.
142. Chatterjee SS, Otto M. 2013. Improved understanding of factors driving methicillin-resistant *Staphylococcus aureus* epidemic waves. *Clin Epidemiol* 5:205-17.
143. Nimmo GR, Coombs GW, Pearson JC, O'Brien FG, Christiansen KJ, Turnidge JD, Gosbell IB, Collignon P, McLaws ML. 2006. Methicillin-resistant *Staphylococcus aureus* in the Australian community: an evolving epidemic. *Med J Aust* 184:384-8.
144. Skov R, Christiansen K, Dancer SJ, Daum RS, Dryden M, Huang YC, Lowy FD. 2012. Update on the prevention and control of community-acquired methicillin-resistant *Staphylococcus aureus* (CA-MRSA). *Int J Antimicrob Agents* 39:193-200.
145. Howden BP, Davies JK, Johnson PDR, Stinear TP, Grayson ML. 2010. Reduced vancomycin susceptibility in *Staphylococcus aureus*, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: resistance mechanisms, laboratory detection, and clinical implications. *Clin Microbiol Rev* 23:99-139.
146. Levine DP. 2006. Vancomycin: a history. *Clin Infect Dis* 42 Suppl 1:S5-12.
147. Crotty MP, Krekel T, Burnham CA, Ritchie DJ. 2016. New Gram-Positive Agents: the Next Generation of Oxazolidinones and Lipoglycopeptides. *J Clin Microbiol* 54:2225-32.
148. Conlon BP, Rowe SE, Gandt AB, Nuxoll AS, Donegan NP, Zalis EA, Clair G, Adkins JN, Cheung AL, Lewis K. 2016. Persister formation in *Staphylococcus aureus* is associated with ATP depletion. *Nat Microbiol* 1.
149. Arthur M, Molinas C, Depardieu F, Courvalin P. 1993. Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *J Bacteriol* 175:117-27.
150. Courvalin P. 2006. Vancomycin resistance in gram-positive cocci. *Clin Infect Dis* 42 Suppl 1:S25-34.
151. Chang S, Sievert DM, Hageman JC, Boulton ML, Tenover FC, Downes FP, Shah S, Rudrik JT, Pupp GR, Brown WJ, Cardo D, Fridkin SK, Vancomycin-Resistant *Staphylococcus aureus* Investigative T. 2003. Infection with vancomycin-resistant *Staphylococcus aureus* containing the vanA resistance gene. *N Engl J Med* 348:1342-7.
152. Perichon B, Courvalin P. 2009. VanA-type vancomycin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 25:1580-7.
153. Gardete S, Tomasz A. 2014. Mechanisms of vancomycin resistance in *Staphylococcus aureus*. *J Clin Invest* 124:2836-40.
154. Howden BP, Peleg AY, Stinear TP. 2014. The evolution of vancomycin intermediate *Staphylococcus aureus* (VISA) and heterogenous-VISA. *Infect Genet Evol* 21:575-582.
155. Gao W, Cameron DR, Davies JK, Kostoulas X, Stepanell J, Tuck KL, Yeaman MR, Peleg AY, Stinear TP, Howden BP. 2013. The rpoB_{H481Y} rifampicin resistance mutation and an active stringent response reduce virulence and increase resistance to innate immune responses in *Staphylococcus aureus*. *J Infect Dis* 207:929-39.

156. Katayama Y, Sekine M, Hishinuma T, Aiba Y, Hiramatsu K. 2016. Complete Reconstitution of the Vancomycin-Intermediate *Staphylococcus aureus* Phenotype of Strain Mu50 in Vancomycin-Susceptible *S. aureus*. *Antimicrob Agents Chemother* 60:3730-42.
157. van Hal SJ, Paterson DL. 2011. Systematic review and meta-analysis of the significance of heterogeneous vancomycin-intermediate *Staphylococcus aureus* isolates. *Antimicrob Agents Chemother* 55:405-10.
158. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MVW, Anderson TL, Roberts SA, Warren SJC, Coombs GW, Tan H-L, Gao W, Johnson PDR, Howden BP. 2014. Genetic and molecular predictors of high vancomycin minimum inhibitory concentration in *Staphylococcus aureus* bacteremia isolates. *J Clin Micro* doi:10.1128/JCM.01320-10.
159. Arbeit RD, Maki D, Tally FP, Campanaro E, Eisenstein BI, Daptomycin, Investigators. 2004. The safety and efficacy of daptomycin for the treatment of complicated skin and skin-structure infections. *Clin Infect Dis* 38:1673-81.
160. Bayer AS, Schneider T, Sahl HG. 2013. Mechanisms of daptomycin resistance in *Staphylococcus aureus*: role of the cell membrane and cell wall. *Ann N Y Acad Sci* 1277:139-58.
161. Steenbergen JN, Alder J, Thorne GM, Tally FP. 2005. Daptomycin: a lipopeptide antibiotic for the treatment of serious Gram-positive infections. *J Antimicrob Chemother* 55:283-8.
162. Miller WR, Bayer AS, Arias CA. 2016. Mechanism of Action and Resistance to Daptomycin in *Staphylococcus aureus* and Enterococci. *Cold Spring Harb Perspect Med* 6.
163. Ernst CM, Peschel A. 2011. Broad-spectrum antimicrobial peptide resistance by MprF-mediated aminoacylation and flipping of phospholipids. *Mol Microbiol* 80:290-9.
164. Ernst CM, Staubitz P, Mishra NN, Yang SJ, Hornig G, Kalbacher H, Bayer AS, Kraus D, Peschel A. 2009. The bacterial defensin resistance protein MprF consists of separable domains for lipid lysinylation and antimicrobial peptide repulsion. *PLoS Pathog* 5:e1000660.
165. Howden BP, McEvoy CRE, Allen DL, Chua K, Gao W, Harrison PF, Bell J, Coombs G, Bennett-Wood V, Porter JL, Robins-Browne R, Davies JK, Seemann T, Stinear TP. 2011. Evolution of multidrug resistance during *Staphylococcus aureus* infection involves mutation of the essential two component regulator WalKR. *PLoS Pathog* 7:e1002359.
166. Cui L, Tominaga E, Neoh HM, Hiramatsu K. 2006. Correlation between Reduced Daptomycin Susceptibility and Vancomycin Resistance in Vancomycin-Intermediate *Staphylococcus aureus*. *Antimicrob Agents Chemother* 50:1079-82.
167. Peleg AY, Miyakis S, Ward DV, Earl AM, Rubio A, Cameron DR, Pillai S, Moellering RC, Jr., Eliopoulos GM. 2012. Whole genome characterization of the mechanisms of daptomycin resistance in clinical and laboratory derived isolates of *Staphylococcus aureus*. *PLoS One* 7:e28316.
168. Berti AD, Baines SL, Howden BP, Sakoulas G, Nizet V, Proctor RA, Rose WE. 2015. Heterogeneity of genetic pathways toward daptomycin nonsusceptibility in *Staphylococcus aureus* determined by adjunctive antibiotics. *Antimicrob Agents Chemother* 59:2799-806.
169. Jones T, Yeaman MR, Sakoulas G, Yang SJ, Proctor RA, Sahl HG, Schrenzel J, Xiong YQ, Bayer AS. 2008. Failures in clinical treatment of *Staphylococcus aureus* Infection with daptomycin are associated with alterations in surface charge, membrane phospholipid asymmetry, and drug binding. *Antimicrob Agents Chemother* 52:269-78.

170. Kelley PG, Gao W, Ward PB, Howden BP. 2011. Daptomycin non-susceptibility in vancomycin-intermediate *Staphylococcus aureus* (VISA) and heterogeneous-VISA (hVISA): implication for therapy after vancomycin treatment failure. *J Antimicrob Chemother* 66:1057-1060.
171. Patel JB, Jevitt LA, Hageman J, McDonald LC, Tenover FC. 2006. An association between reduced susceptibility to daptomycin and reduced susceptibility to vancomycin in *Staphylococcus aureus*. *Clin Infect Dis* 42:1652-3.
172. Sader HS, Fritsche TR, Jones RN. 2006. Daptomycin bactericidal activity and correlation between disk and broth microdilution method results in testing of *Staphylococcus aureus* strains with decreased susceptibility to vancomycin. *Antimicrob Agents Chemother* 50:2330-6.
173. Gupta S, Bhatia G, Sharma A, Saxena S. 2018. Host defense peptides: An insight into the antimicrobial world. *J Oral Maxillofac Pathol* 22:239-244.
174. Yang D, Biragyn A, Hoover DM, Lubkowski J, Oppenheim JJ. 2004. Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense. *Annu Rev Immunol* 22:181-215.
175. Mishra NN, McKinnell J, Yeaman MR, Rubio A, Nast CC, Chen L, Kreiswirth BN, Bayer AS. 2011. *In vitro* cross-resistance to daptomycin and host defense cationic antimicrobial peptides in clinical methicillin-resistant *Staphylococcus aureus* isolates. *Antimicrob Agents Chemother* 55:4012-8.
176. Findlay F, Proudfoot L, Stevens C, Barlow PG. 2016. Cationic host defense peptides; novel antimicrobial therapeutics against Category A pathogens and emerging infections. *Pathog Glob Health* 110:137-47.
177. Fleitas O, Franco OL. 2016. Induced Bacterial Cross-Resistance toward Host Antimicrobial Peptides: A Worrying Phenomenon. *Front Microbiol* 7:381.
178. Williamson DA, Carter GP, Howden BP. 2017. Current and Emerging Topical Antibacterials and Antiseptics: Agents, Action, and Resistance Patterns. *Clin Microbiol Rev* 30:827-860.
179. Koning S, van der Sande R, Verhagen AP, van Suijlekom-Smit LW, Morris AD, Butler CC, Berger M, van der Wouden JC. 2012. Interventions for impetigo. *Cochrane Database Syst Rev* 1:CD003261.
180. Horner C, Mawer D, Wilcox M. 2012. Reduced susceptibility to chlorhexidine in staphylococci: is it increasing and does it matter? *J Antimicrob Chemother* 67:2547-59.
181. Milstone AM, Passaretti CL, Perl TM. 2008. Chlorhexidine: expanding the armamentarium for infection control and prevention. *Clin Infect Dis* 46:274-81.
182. Aly R, Maibach HI. 1988. Comparative antibacterial efficacy of a 2-minute surgical scrub with chlorhexidine gluconate, povidone-iodine, and chloroxylenol sponge-brushes. *Am J Infect Control* 16:173-7.
183. Rosenberg A, Alatary SD, Peterson AF. 1976. Safety and efficacy of the antiseptic chlorhexidine gluconate. *Surg Gynecol Obstet* 143:789-92.
184. Godtfredsen W, Roholt K, Tybring L. 1962. Fucidin: a new orally active antibiotic. *Lancet* 1:928-31.
185. Dobie D, Gray J. 2004. Fusidic acid resistance in *Staphylococcus aureus*. *Arch Dis Child* 89:74-7.
186. Verbist L. 1990. The antimicrobial activity of fusidic acid. *J Antimicrob Chemother* 25 Suppl B:1-5.
187. Borg A, Holm M, Shiroyama I, Hauryliuk V, Pavlov M, Sanyal S, Ehrenberg M. 2015. Fusidic acid targets elongation factor G in several stages of translocation on the bacterial ribosome. *J Biol Chem* 290:3440-54.

188. Nagaev I, Bjorkman J, Andersson DI, Hughes D. 2001. Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. *Mol Microbiol* 40:433-9.
189. Besier S, Ludwig A, Brade V, Wichelhaus TA. 2003. Molecular analysis of fusidic acid resistance in *Staphylococcus aureus*. *Mol Microbiol* 47:463-9.
190. Laurberg M, Kristensen O, Martemyanov K, Gudkov AT, Nagaev I, Hughes D, Liljas A. 2000. Structure of a mutant EF-G reveals domain III and possibly the fusidic acid binding site. *J Mol Biol* 303:593-603.
191. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V. 2009. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* 326:694-9.
192. Cox G, Thompson GS, Jenkins HT, Peske F, Savelsbergh A, Rodnina MV, Wintermeyer W, Homans SW, Edwards TA, O'Neill AJ. 2012. Ribosome clearance by FusB-type proteins mediates resistance to the antibiotic fusidic acid. *Proc Natl Acad Sci U S A* 109:2102-7.
193. Ellington MJ, Reuter S, Harris SR, Holden MT, Cartwright EJ, Greaves D, Gerver SM, Hope R, Brown NM, Torok ME, Parkhill J, Koser CU, Peacock SJ. 2015. Emergent and evolving antimicrobial resistance cassettes in community-associated fusidic acid and meticillin-resistant *Staphylococcus aureus*. *Int J Antimicrob Agents* 45:477-84.
194. Baines SL, Howden BP, Heffernan H, Stinear TP, Carter GP, Seemann T, Kwong JC, Ritchie SR, Williamson DA. 2016. Rapid Emergence and Evolution of *Staphylococcus aureus* Clones Harboring *fusC*-Containing Staphylococcal Cassette Chromosome Elements. *Antimicrob Agents Chemother* 60:2359-65.
195. Gilbert P, Moore LE. 2005. Cationic antiseptics: diversity of action under a common epithet. *J Appl Microbiol* 99:703-15.
196. Russell AD. 2002. Introduction of biocides into clinical practice and the impact on antibiotic-resistant bacteria. *Symp Ser Soc Appl Microbiol*:121S-135S.
197. Mitchell BA, Brown MH, Skurray RA. 1998. QacA multidrug efflux pump from *Staphylococcus aureus*: comparative analysis of resistance to diamidines, biguanidines, and guanylhya zones. *Antimicrob Agents Chemother* 42:475-7.
198. Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA. 1996. Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity. *Proc Natl Acad Sci U S A* 93:3630-5.
199. Hardy K, Sunnucks K, Gil H, Shabir S, Trampari E, Hawkey P, Webber M. 2018. Increased Usage of Antiseptics Is Associated with Reduced Susceptibility in Clinical Isolates of *Staphylococcus aureus*. *MBio* 9.
200. DeMarco CE, Cushing LA, Frempong-Manso E, Seo SM, Jaravaza TA, Kaatz GW. 2007. Efflux-related resistance to norfloxacin, dyes, and biocides in bloodstream isolates of *Staphylococcus aureus*. *Antimicrob Agents Chemother* 51:3235-9.
201. Kaatz GW, Thyagarajan, R. V., Seo, S. M. 2005. Effect of promoter region mutations and *mgrA* overexpression on transcription of *norA*, which encodes a *Staphylococcus aureus* multidrug efflux transporter. *Antimicrob Agents Chemother* 49:161-169.
202. Kaatz GW, Seo SM, Ruble CA. 1993. Efflux-mediated fluoroquinolone resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 37:1086-94.

203. Neyfakh AA, Borsch CM, Kaatz GW. 1993. Fluoroquinolone resistance protein NorA of *Staphylococcus aureus* is a multidrug efflux transporter. *Antimicrob Agents Chemother* 37:128-9.
204. Baines SL, Jensen SO, Firth N, Goncalves da Silva A, Seemann T, Carter GP, Williamson DA, Howden BP, Stinear TP. 2019. Remodeling of pSK1 Family Plasmids and Enhanced Chlorhexidine Tolerance in a Dominant Hospital Lineage of Methicillin-Resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 63.
205. Rush MG, Gordon, C.N., Novick, R. P., Warner, R. C. 1969. Penicillinase plasmid DNA from *Staphylococcus aureus*. *Proc Natl Acad Sci USA* 63:1304-1310.
206. Croes S, Beisser PS, Terporten PH, Neef C, Deurenberg RH, Stobberingh EE. 2010. Diminished in vitro antibacterial activity of oxacillin against clinical isolates of borderline oxacillin-resistant *Staphylococcus aureus*. *Clin Microbiol Infect* 16:979-85.
207. Tomasz A, Drugeon HB, de Lencastre HM, Jabes D, McDougall L, Bille J. 1989. New mechanism for methicillin resistance in *Staphylococcus aureus*: clinical isolates that lack the PBP 2a gene and contain normal penicillin-binding proteins with modified penicillin-binding capacity. *Antimicrob Agents Chemother* 33:1869-74.
208. Long SW, Olsen RJ, Mehta SC, Palzkill T, Cernoch PL, Perez KK, Musick WL, Rosato AE, Musser JM. 2014. PBP2a mutations causing high-level Ceftaroline resistance in clinical methicillin-resistant *Staphylococcus aureus* isolates. *Antimicrob Agents Chemother* 58:6668-74.
209. Chan LC, Basuino L, Diep B, Hamilton S, Chatterjee SS, Chambers HF. 2015. Ceftobiprole- and ceftaroline-resistant methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 59:2960-3.
210. Melo-Cristino J, Resina C, Manuel V, Lito L, Ramirez M. 2013. First case of infection with vancomycin-resistant *Staphylococcus aureus* in Europe. *Lancet* 382:205.
211. Werth BJ, Jain R, Hahn A, Cummings L, Weaver T, Waalkes A, Sengupta D, Salipante SJ, Rakita RM, Butler-Wu SM. 2018. Emergence of dalbavancin non-susceptible, vancomycin-intermediate *Staphylococcus aureus* (VISA) after treatment of MRSA central line-associated bloodstream infection with a dalbavancin- and vancomycin-containing regimen. *Clin Microbiol Infect* 24:429 e1-429 e5.
212. Kussmann M, Karer M, Obermueller M, Schmidt K, Barousch W, Moser D, Nehr M, Ramharter M, Poepl W, Makristathis A, Winkler S, Thalhammer F, Burgmann H, Lagler H. 2018. Emergence of a dalbavancin induced glycopeptide/lipoglycopeptide non-susceptible *Staphylococcus aureus* during treatment of a cardiac device-related endocarditis. *Emerg Microbes Infect* 7:202.
213. Friedman L, Alder, J. D., Silverman, J. A. 2006. Genetic changes that correlate with reduced susceptibility to Daptomycin in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 50:2137-2145.
214. Ramirez MS, Tolmasky ME. 2010. Aminoglycoside modifying enzymes. *Drug Resist Updat* 13:151-71.
215. Rouch DA, Byrne, M. E., Kong, Y. C., Skurray, R. A. 1987. The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and nucleotide sequence analysis. *J Gen Microbiol* 133:3039-3052.
216. Holden MTG, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, Batra R, Parkhill J, Bentley SD, Edgeworth JD. 2010. Genome sequence of a recently emerged, highly transmissible, multi-

- antibiotic- and antiseptic-resistant variant of methicillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). J Bacteriol 192:888-892.
217. McKenzie T, Hoshino T, Tanaka T, Sueoka N. 1986. The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. Plasmid 15:93-103.
 218. McAleese F, Petersen P, Ruzin A, Dunman PM, Murphy E, Projan SJ, Bradford PA. 2005. A novel MATE family efflux pump contributes to the reduced susceptibility of laboratory-derived *Staphylococcus aureus* mutants to tigecycline. Antimicrob Agents Chemother 49:1865-71.
 219. Weisblum B. 1995. Erythromycin resistance by ribosome modification. Antimicrob Agents Chemother 39:577-85.
 220. Sharkey LK, Edwards TA, O'Neill AJ. 2016. ABC-F Proteins Mediate Antibiotic Resistance through Ribosomal Protection. MBio 7:e01975.
 221. Kehrenberg C, Schwarz S, Jacobsen L, Hansen LH, Vester B. 2005. A new mechanism for chloramphenicol, florfenicol and clindamycin resistance: methylation of 23S ribosomal RNA at A2503. Mol Microbiol 57:1064-73.
 222. Long KS, Vester B. 2012. Resistance to linezolid caused by modifications at its binding site on the ribosome. Antimicrob Agents Chemother 56:603-12.
 223. Prunier AL, Malbruny B, Laurans M, Brouard J, Duhamel JF, Leclercq R. 2003. High rate of macrolide resistance in *Staphylococcus aureus* strains from patients with cystic fibrosis reveals high proportions of hypermutable strains. J Infect Dis 187:1709-16.
 224. O'Neill AJ, Chopra I. 2006. Molecular basis of fusB-mediated resistance to fusidic acid in *Staphylococcus aureus*. Mol Microbiol 59:664-76.
 225. O'Brien FG, Price C, Grubb WB, Gustafson JE. 2002. Genetic characterization of the fusidic acid and cadmium resistance determinants of *Staphylococcus aureus* plasmid pUB101. J Antimicrob Chemother 50:313-21.
 226. Thomas CM, Hotherhall J, Willis CL, Simpson TJ. 2010. Resistance to and synthesis of the antibiotic mupirocin. Nat Rev Microbiol 8:281-9.
 227. Khoshnood S, Heidary M, Asadi A, Soleimani S, Motahar M, Savari M, Saki M, Abdi M. 2019. A review on mechanism of action, resistance, synergism, and clinical implications of mupirocin against *Staphylococcus aureus*. Biomed Pharmacother 109:1809-1818.
 228. Carter GP, Schultz MB, Baines SL, Goncalves da Silva A, Heffernan H, Tiong A, Pham PH, Monk IR, Stinear TP, Howden BP, Williamson DA. 2018. Topical Antibiotic Use Coselects for the Carriage of Mobile Genetic Elements Conferring Resistance to Unrelated Antimicrobials in *Staphylococcus aureus*. Antimicrob Agents Chemother 62.
 229. Dale GE, Broger C, D'Arcy A, Hartman PG, DeHoogt R, Jolidon S, Kompis I, Labhardt AM, Langen H, Locher H, Page MG, Stuber D, Then RL, Wipf B, Oefner C. 1997. A single amino acid substitution in *Staphylococcus aureus* dihydrofolate reductase determines trimethoprim resistance. J Mol Biol 266:23-30.
 230. Rouch DA, Messerotti LJ, Loo LS, Jackson CA, Skurray RA. 1989. Trimethoprim resistance transposon Tn4003 from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS257. Mol Microbiol 3:161-75.

231. Kadlec K, Schwarz S. 2009. Identification of a novel trimethoprim resistance gene, *dfrK*, in a methicillin-resistant *Staphylococcus aureus* ST398 strain and its physical linkage to the tetracycline resistance gene *tet(L)*. *Antimicrob Agents Chemother* 53:776-8.
232. Sekiguchi J, Tharavichitkul P, Miyoshi-Akiyama T, Chupia V, Fujino T, Araake M, Irie A, Morita K, Kuratsuji T, Kirikae T. 2005. Cloning and characterization of a novel trimethoprim-resistant dihydrofolate reductase from a nosocomial isolate of *Staphylococcus aureus* CM.S2 (IMCJ1454). *Antimicrob Agents Chemother* 49:3948-51.
233. Hampele IC, D'Arcy A, Dale GE, Kostrewa D, Nielsen J, Oefner C, Page MG, Schonfeld HJ, Stuber D, Then RL. 1997. Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. *J Mol Biol* 268:21-30.
234. Griffith EC, Wallace MJ, Wu Y, Kumar G, Gajewski S, Jackson P, Phelps GA, Zheng Z, Rock CO, Lee RE, White SW. 2018. The Structural and Functional Basis for Recurring Sulfa Drug Resistance Mutations in *Staphylococcus aureus* Dihydropteroate Synthase. *Front Microbiol* 9:1369.
235. Hooper DC, Jacoby GA. 2015. Mechanisms of drug resistance: quinolone resistance. *Ann N Y Acad Sci* 1354:12-31.
236. Hooper DC. 2002. Fluoroquinolone resistance among Gram-positive cocci. *Lancet Infect Dis* 2:530-8.
237. Aubry-Damon H, Soussy CJ, Courvalin P. 1998. Characterization of mutations in the *rpoB* gene that confer rifampin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 42:2590-4.
238. Homma T, Nuxoll A, Gandt AB, Ebner P, Engels I, Schneider T, Gotz F, Lewis K, Conlon BP. 2016. Dual Targeting of Cell Wall Precursors by Teixobactin Leads to Cell Lysis. *Antimicrob Agents Chemother* 60:6510-6517.
239. Balasubramanian D, Harper L, Shopsin B, Torres VJ. 2017. *Staphylococcus aureus* pathogenesis in diverse host environments. *Pathog Dis* 75.
240. Patti JM, Allen BL, McGavin MJ, Hook M. 1994. MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu Rev Microbiol* 48:585-617.
241. Foster TJ, Hook M. 1998. Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol* 6:484-8.
242. Liu GY. 2009. Molecular pathogenesis of *Staphylococcus aureus* infection. *Pediatr Res* 65:71R-77R.
243. Li M, Du X, Villaruz AE, Diep BA, Wang D, Song Y, Tian Y, Hu J, Yu F, Lu Y, Otto M. 2012. MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med* 18:816-819.
244. Bien J, Sokolova O, Bozko P. 2011. Characterization of Virulence Factors of *Staphylococcus aureus*: Novel Function of Known Virulence Factors That Are Implicated in Activation of Airway Epithelial Proinflammatory Response. *J Pathog* 2011:601905.
245. Wanner S, Schade J, Keinhorster D, Weller N, George SE, Kull L, Bauer J, Grau T, Winstel V, Stoy H, Kretschmer D, Kolata J, Wolz C, Broker BM, Weidenmaier C. 2017. Wall teichoic acids mediate increased virulence in *Staphylococcus aureus*. *Nat Microbiol* 2:16257.
246. Thomer L, Schneewind O, Missiakas D. 2016. Pathogenesis of *Staphylococcus aureus* Bloodstream Infections. *Annu Rev Pathol* 11:343-64.
247. O'Riordan K, Lee JC. 2004. *Staphylococcus aureus* capsular polysaccharides. *Clin Microbiol Rev* 17:218-34.

248. Deisenhofer J. 1981. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. *Biochemistry* 20:2361-70.
249. Foster TJ. 2005. Immune evasion by staphylococci. *Nat Rev Microbiol* 3:948-58.
250. Patot S, Imbert PR, Baude J, Martins Simoes P, Campergue JB, Louche A, Nijland R, Bes M, Tristan A, Laurent F, Fischer A, Schrenzel J, Vandenesch F, Salcedo SP, Francois P, Lina G. 2017. The TIR Homologue Lies near Resistance Genes in *Staphylococcus aureus*, Coupling Modulation of Virulence and Antimicrobial Susceptibility. *PLoS Pathog* 13:e1006092.
251. Reyes-Robles T, Torres VJ. 2017. *Staphylococcus aureus* Pore-Forming Toxins. *Curr Top Microbiol Immunol* 409:121-144.
252. Narita S, Kaneko J, Chiba J, Piemont Y, Jarraud S, Etienne J, Kamio Y. 2001. Phage conversion of Panton-Valentine leukocidin in *Staphylococcus aureus*: molecular analysis of a PVL-converting phage, phiSLT. *Gene* 268:195-206.
253. Peschel A, Otto M. 2013. Phenol-soluble modulins and staphylococcal infection. *Nat Rev Microbiol* 11:667-73.
254. Dinges MM, Orwin PM, Schlievert PM. 2000. Exotoxins of *Staphylococcus aureus*. *Clin Microbiol Rev* 13:16-34, table of contents.
255. Torres VJ, Pishchany G, Humayun M, Schneewind O, Skaar EP. 2006. *Staphylococcus aureus* IsdB is a hemoglobin receptor required for heme iron utilization. *J Bacteriol* 188:8421-9.
256. Hammer ND, Skaar EP. 2011. Molecular mechanisms of *Staphylococcus aureus* iron acquisition. *Annu Rev Microbiol* 65:129-47.
257. Archer NK, Mazaitis MJ, Costerton JW, Leid JG, Powers ME, Shirtliff ME. 2011. *Staphylococcus aureus* biofilms: properties, regulation, and roles in human disease. *Virulence* 2:445-59.
258. Lister JL, Horswill AR. 2014. *Staphylococcus aureus* biofilms: recent developments in biofilm dispersal. *Front Cell Infect Microbiol* 4:178.
259. Cao S, Huseby DL, Brandis G, Hughes D. 2017. Alternative Evolutionary Pathways for Drug-Resistant Small Colony Variant Mutants in *Staphylococcus aureus*. *MBio* 8.
260. Proctor RA, van Langevelde P, Kristjansson M, Maslow JN, Arbeit RD. 1995. Persistent and relapsing infections associated with small-colony variants of *Staphylococcus aureus*. *Clin Infect Dis* 20:95-102.
261. Kahl B, Herrmann M, Everding AS, Koch HG, Becker K, Harms E, Proctor RA, Peters G. 1998. Persistent infection with small colony variant strains of *Staphylococcus aureus* in patients with cystic fibrosis. *J Infect Dis* 177:1023-9.
262. von Eiff C, Proctor RA, Peters G. 2000. *Staphylococcus aureus* small colony variants: formation and clinical impact. *Int J Clin Pract Suppl*:44-9.
263. Guerillot R, Kostoulas X, Donovan L, Li L, Carter GP, Hachani A, Vandellannoote K, Giulieri S, Monk IR, Kunimoto M, Starrs L, Burgio G, Seemann T, Peleg AY, Stinear TP, Howden BP. 2019. Unstable chromosome rearrangements in *Staphylococcus aureus* cause phenotype switching associated with persistent infections. *Proc Natl Acad Sci U S A* 116:20135-20140.
264. Diep BA, Gill S. R., Chang R. F., Phan T. H., Chen J. H., Davidson M. G., Lin F., Lin J., Carleton H. A., Mongodin E. F., Sensabaugh G. F., Perdreau-Remington F. 2006. Complete genome sequence

- of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. Lancet 367:731-739.
265. Montgomery CP, Boyle-Vavra S, Daum RS. 2009. The arginine catabolic mobile element is not associated with enhanced virulence in experimental invasive disease caused by the community-associated methicillin-resistant *Staphylococcus aureus* USA300 genetic background. Infect Immun 77:2650-6.
 266. Diep BA, Stone GG, Li B, Graber C, J., Miller A, des Etages S-A, Jones A, Palazzolo-Ballance AM, Perdreau-Remington F, Sensabaugh GF, DeLeo FR, Chambers HF. 2008. The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus*. J Infect Dis 197.
 267. Urushibara N, Kawaguchiya M, Kobayashi N. 2012. Two novel arginine catabolic mobile elements and staphylococcal chromosome cassette *mec* composite islands in community-acquired methicillin-resistant *Staphylococcus aureus* genotypes ST5-MRSA-V and ST5-MRSA-II. J Antimicrob Chemother 67:1828-34.
 268. Cunin R, Glansdorff N, Pierard A, Stalon V. 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiol Rev 50:314-52.
 269. Recsei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A, Novick RP. 1986. Regulation of exoprotein gene expression in *Staphylococcus aureus* by agar. Mol Gen Genet 202:58-61.
 270. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. 1993. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12:3967-75.
 271. Novick RP, Geisinger E. 2008. Quorum sensing in staphylococci. Annu Rev Genet 42:541-64.
 272. Wang B, Muir TW. 2016. Regulation of Virulence in *Staphylococcus aureus*: Molecular Mechanisms and Remaining Puzzles. Cell Chem Biol 23:214-224.
 273. Herbert S, Ziebandt A-K, Ohlsen K, Schafer T, Hecker M, Albrecht D, Novick RP, Gotz F. 2010. Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. Infect Immun 78:2877-2889.
 274. Altman DR, Sullivan MJ, Chacko KI, Balasubramanian D, Pak TR, Sause WE, Kumar K, Sebra R, Deikus G, Attie O, Rose H, Lewis M, Fulmer Y, Bashir A, Kasarskis A, Schadt EE, Richardson AR, Torres VJ, Shopsin B, van Bakel H. 2018. Genome Plasticity of *agr*-Defective *Staphylococcus aureus* during Clinical Infection. Infect Immun 86.
 275. Shopsin B, Eaton C, Wasserman GA, Mathema B, Adhikari RP, Agolory S, Altman DR, Holzman RS, Kreiswirth BN, Novick RP. 2010. Mutations in *agr* do not persist in natural populations of methicillin-resistant *Staphylococcus aureus*. J Infect Dis 202:1593-9.
 276. Chong YP, Kim ES, Park SJ, Park KH, Kim T, Kim MN, Kim SH, Lee SO, Choi SH, Woo JH, Jeong JY, Kim YS. 2013. Accessory gene regulator (*agr*) dysfunction in *Staphylococcus aureus* bloodstream isolates from South Korean patients. Antimicrob Agents Chemother 57:1509-12.
 277. Hodille E, Rose W, Diep BA, Goutelle S, Lina G, Dumitrescu O. 2017. The Role of Antibiotics in Modulating Virulence in *Staphylococcus aureus*. Clin Microbiol Rev 30:887-917.

278. Painter KL, Krishna A, Wigneshweraraj S, Edwards AM. 2014. What role does the quorum-sensing accessory gene regulator system play during *Staphylococcus aureus* bacteremia? Trends Microbiol 22:676-85.
279. Vandecraen J, Chandler M, Aertsen A, Van Houdt R. 2017. The impact of insertion sequences on bacterial genome plasticity and adaptability. Crit Rev Microbiol 43:709-730.
280. Firth N, Skurray RA. 1998. Mobile elements in the evolution and spread of multiple-drug resistance in staphylococci. Drug Resist Updat 1:49-58.
281. Katayama Y, Ito, T., Hiramatsu, K. 2000. A new class of genetic element, Staphylococcus Cassette Chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. Antimicrob Agents Chemother 44:1549-1555.
282. Ito T, Ma XX, Takeuchi F, Okuma K, Yuzawa H, Hiramatsu K. 2004. Novel type V staphylococcal cassette chromosome *mec* driven by a novel cassette chromosome recombinase, *ccrC*. Antimicrob Agents Chemother 48:2637-51.
283. International Working Group on the Classification of Staphylococcal Cassette Chromosome E. 2009. Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. Antimicrob Agents Chemother 53:4961-7.
284. Baig S, Johannesen TB, Overballe-Petersen S, Larsen J, Larsen AR, Stegger M. 2018. Novel SCC*mec* type XIII (9A) identified in an ST152 methicillin-resistant *Staphylococcus aureus*. Infect Genet Evol 61:74-76.
285. Heusser R, Ender M, Berger-Bachi B, McCallum N. 2007. Mosaic staphylococcal cassette chromosome *mec* containing two recombinase loci and a new *mec* complex, B2. Antimicrob Agents Chemother 51:390-3.
286. Boyle-Vavra S, Ereshefsky B, Wang CC, Daum RS. 2005. Successful multiresistant community-associated methicillin-resistant *Staphylococcus aureus* lineage from Taipei, Taiwan, that carries either the novel Staphylococcal chromosome cassette *mec* (SCC*mec*) type VT or SCC*mec* type IV. J Clin Microbiol 43:4719-30.
287. Novick RP. 1989. Staphylococcal plasmids and their replication. Annu Rev Microbiol 43:537-65.
288. Jensen SO, Apisiridej S, Kwong SM, Yang YH, Skurray RA, Firth N. 2010. Analysis of the prototypical *Staphylococcus aureus* multiresistance plasmid pSK1. Plasmid 64:135-142.
289. Gillespie MT, Lyon BR, Messerotti LJ, Skurray RA. 1987. Chromosome- and plasmid-mediated gentamicin resistance in *Staphylococcus aureus* encoded by Tn4001. J Med Microbiol 24:139-44.
290. Apisiridej S, Leelaporn A, Scaramuzzi CD, Skurray RA, Firth N. 1997. Molecular analysis of a mobilizable theta-mode trimethoprim resistance plasmid from coagulase-negative staphylococci. Plasmid 38:13-24.
291. Rouch DA, Cram DS, DiBerardino D, Littlejohn TG, Skurray RA. 1990. Efflux-mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins. Mol Microbiol 4:2051-62.
292. Lyon BR, May JW, Skurray RA. 1984. Tn4001: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. Mol Gen Genet 193:554-6.

293. Rouch DA, Skurray RA. 1989. IS257 from *Staphylococcus aureus*: member of an insertion sequence superfamily prevalent among gram-positive and gram-negative bacteria. *Gene* 76:195-205.
294. Koser CU, Holder, M. T. G., Ellington, M. J., Cartwright, E. J. P., Brown, N. M., Ogilvey-Stuart, A. L., Hsu, L. Y., Chewapreecha, C., Croucher, N. J., Harris, S. R., Sanders, M., Enright, M. C., Dougan, G., Bentley, S. D., Parkhill, J., Fraser, L. J., Betley, J. R., Schulz-Trieglaff, O. B., Smith, G. P. & Peacock, S. J. 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 366:2267-2275.
295. Baker S, Thomson N, Weill FX, Holt KE. 2018. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. *Science* 360:733-738.
296. Pavillard R, Harvey K, Douglas D, Hewstone A, Andrew J, Collopy B, Asche V, Carson P, Davidson A, Gilbert G, Spicer J, Tosoloni F. 1982. Epidemic of hospital-acquired infection due to methicillin-resistant *Staphylococcus aureus* in major Victorian hospitals. *Med J Aust* 29:451-454.
297. Cookson BD, Philips I. 1988. Epidemic methicillin-resistant *Staphylococcus aureus*. *J Antimicrob Chemother* 21 (Suppl.C):57-65.
298. Dubin DT, Chikramane SG, Inglis B, Matthews PR, Stewart PR. 1992. Physical mapping of the *mec* region of an Australian methicillin-resistant *Staphylococcus aureus* lineage and a closely related American strain. *J Gen Microbiol* 138:169-180.
299. Gray RR, Tatem AJ, Johnson JA, Alekseyenko AV, Pybus OG, Suchard MA, Salemi M. 2011. Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a bayesian framework. *Mol Biol Evol* 28:1593-1603.
300. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327:469-474.
301. Castillo-Ramirez S, Corander J, Marttinen P, Aldeljawi M, Hanage WP, Westh H, Boye K, Bentley SD, Parkhill J, Holden MT, Feil EJ. 2012. Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol* 13:R126.
302. Smyth DS, McDougal LK, Gran FW, Manoharan A, Enright MC, Song J-H, de Lencastre H, Robinson DA. 2009. Population structure of a hybrid clone of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. *PLoS ONE* 5:e8582.
303. Hsu LY, Harris SR, Chlebowicz MA, Lindsay JA, Koh TH, Krishnan P, Tan TY, Hon PY, Grubb WB, Bentley SD, Parkhill J, Peacock SJ, Holden MT. 2015. Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol* 16:81.
304. Feil EJ, Nickerson EK, Chantratita N, Wuthiekanun V, Srisomang P, Cousins R, Pan W, Zhang G, Xu B, Day NPJ, Peacock SJ. 2008. Rapid detection of the pandemic methicillin-resistant *Staphylococcus aureus* clone ST239, a dominant strain in Asian hospitals. *J Clin Micro* 46:1520-1522.
305. Howden BP, Johnson PDR, Ward PB, Stinear TP, Davies JK. 2006. Isolates with low-level vancomycin resistance associated with persistent methicillin-resistant *Staphylococcus aureus* bacteremia. *Antimicrob Agents Chemother* 50:3039-3047.
306. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MVN, Anderson TL, Roberts SA, Gao W, Christiansen KJ, Coombs GW, Johnson PDR, Howden BP. 2011. Antibiotic choice

- may not explain poorer outcomes in patients with *Staphylococcus aureus* bacteremia and high vancomycin minimum inhibitory concentration. J Infect Dis 204:340-347.
307. David M, Dzamba M, D L, Ilie L, Brudno M. 2009. SHRiMP2: sensitive yet practical short read mapping. Bioinform 25:1966-7.
 308. Stinear TP, Holt KE, Chua KYL, Stepnell J, Tuck KL, Coombs GW, Harrison PF, Seemann T, Howden BP. 2014. Adaptive change inferred from genomic population analysis of the ST93 epidemic clone of community-associated methicillin-resistant *Staphylococcus aureus*. Genome Biol Evol 6:366-378.
 309. Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linder M, McGee L, von Gottberg A, Song J-H, Ko KS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A, Klugman KP, Parkhill J, Hanage WP, Bentley SD. 2011. Rapid pneumococcal evolution in response to clinical intervention. Science 331:430-434.
 310. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinform 22:2688-2690.
 311. Yang Z. 2007. PAML4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591.
 312. Guindon S, Delsuc F, Dufayard JF, Gascuel O. 2009. Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol 537:113-137.
 313. Rambaut A. 2010. Path-O-Gen v1.3: temporal signal investigation tool. <http://tree.bio.ed.ac.uk/software/pathogen/>. Accessed
 314. Lechner M, Findei S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: Detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12.
 315. CLSI. 2012. Performance standards for antimicrobial susceptibility testing: twenty-second informational supplement, vol CLSI document M100-S22. Clinical Laboratory Standards Institute, Wayne, PA.
 316. Walsh TR, Bolmstrom A, Qwarnstrom A, Ho P, Wooton M, Howe RA, MacGowan AP, Diekema D. 2001. Evaluation of current methods for detection of staphylococci with reduced susceptibility to glycopeptides. J Clin Micro 39:2439-2444.
 317. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575.
 318. Lam MMC, Seemann T, Tobias NJ, Chen H, Haring V, Moore RJ, Ballard S, Grayson LM, Johnson PDR, Howden BP, Stinear TP. 2013. Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. BMC Genomics 14.
 319. R Core Team. 2013. R v2.15.3: A language and environment for statistical computing., R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
 320. GraphPad Software. 2014. GraphPad Prism v6.04, GraphPad Software, Inc., San Diego, California USA. <http://www.graphpad.com/>.
 321. The Inkscape Team. 2014. Inkscape v0.48.5. <http://www.inkscape.org/>. Accessed
 322. Costa MO, Beltrame CO, Ferreira FA, Botelho AM, Lima NC, Souza RC, de Almeida LG, Vasconcelos AT, Nicolas MF, Figueiredo AM. 2013. Complete Genome Sequence of a Variant of the Methicillin-Resistant *Staphylococcus aureus* ST239 Lineage, Strain BMB9393, Displaying Superior Ability To Accumulate ica-Independent Biofilm. Genome Announc 1.

323. Botelho AM, Costa MO, Beltrame CO, Ferreira FA, Cortes MF, Bandeira PT, Lima NC, Souza RC, Almeida LG, Vasconcelos AT, Nicolas MF, Figueiredo AM. 2016. Complete genome sequence of an agr-dysfunctional variant of the ST239 lineage of the methicillin-resistant *Staphylococcus aureus* strain GV69 from Brazil. *Stand Genomic Sci* 11:34.
324. Botelho AM, Costa MO, Beltrame CO, Ferreira FA, Lima NC, Costa BS, de Moraes GL, Souza RC, Almeida LG, Vasconcelos AT, Nicolas MF, Figueiredo AM. 2016. Complete Genome Sequence of the MRSA Isolate HC1335 from ST239 Lineage Displaying a Truncated AgrC Histidine Kinase Receptor. *Genome Biol Evol* 8:3187-3192.
325. Howden BP, Seemann T, Harrison PF, McEvoy CR, Stanton J-AL, Rand CJ, Mason CW, Jensen SO, Firth N, Davies JK, Johnson PDR, Stinear TP. 2010. Complete genome sequence of *Staphylococcus aureus* strain JKD6008, an ST239 clones of methicillin-resistant *Staphylococcus aureus* with intermediate level vancomycin resistance. *J Bacteriol* 192:5848-5849.
326. McClure JA, Zhang K. 2017. Complete Genome Sequence of the Methicillin-Resistant *Staphylococcus aureus* Colonizing Strain M92. *Genome Announc* 5.
327. Li Y, Cao B, Zhang Y, Zhou J, Yang B, Wang L. 2011. Complete genome sequence of *Staphylococcus aureus* T0131, an ST239-MRSA-SCCmec type III clones isolated in China. *J Bacteriol* 193:3411-3412.
328. Zhang X, Xu X, Yuan W, Hu Q, Shang W, Hu X, Tong Y, Rao X. 2014. Complete Genome Sequence of *Staphylococcus aureus* XN108, an ST239-MRSA-SCCmec III Strain with Intermediate Vancomycin Resistance Isolated in Mainland China. *Genome Announc* 2.
329. Chen F-J, Lauderdale T-L, Wang L-S, Hsuang I-W. 2013. Complete genome sequence of *Staphylococcus aureus* Z172, a vancomycin-intermediate and daptomycin-nonsusceptible methicillin-resistant strain isolated in Taiwan. *Genome Announc* 1:e01011-13.
330. Kong Z, Zhao P, Liu H, Yu X, Qin Y, Su Z, Wang S, Xu H, Chen J. 2016. Whole-Genome Sequencing for the Investigation of a Hospital Outbreak of MRSA in China. *PLoS One* 11:e0149844.
331. Moradigaravand D, Jamroz D, Mostowy R, Anderson A, Nickerson EK, Thaipadungpanit J, Wuthiekanun V, Limmathurotsakul D, Tandhavanant S, Wikraiphat C, Wongsuvan G, Teerawattanasook N, Jutrakul Y, Srisurat N, Chaimanee P, Eoin West T, Blane B, Parkhill J, Chantratita N, Peacock SJ. 2017. Evolution of the *Staphylococcus argenteus* ST2250 Clone in Northeastern Thailand Is Linked with the Acquisition of Livestock-Associated Staphylococcal Genes. *MBio* 8.
332. Reuter S, Torok ME, Holden MT, Reynolds R, Raven KE, Blane B, Donker T, Bentley SD, Aanensen DM, Grundmann H, Feil EJ, Spratt BG, Parkhill J, Peacock SJ. 2016. Building a genomic framework for prospective MRSA surveillance in the United Kingdom and the Republic of Ireland. *Genome Res* 26:263-70.
333. Alam MT, Petit RA, 3rd, Crispell EK, Thornton TA, Conneely KN, Jiang Y, Satola SW, Read TD. 2014. Dissecting vancomycin-intermediate resistance in *Staphylococcus aureus* using genome-wide association. *Genome Biol Evol* 6:1174-85.
334. Senghore M, Bayliss SC, Kwambana-Adams BA, Foster-Nyarko E, Manneh J, Dione M, Badji H, Ebruke C, Doughty EL, Thorpe HA, Jasinska AJ, Schmitt CA, Cramer JD, Turner TR, Weinstock G, Freimer NB, Pallen MJ, Feil EJ, Antonio M. 2016. Transmission of *Staphylococcus aureus* from Humans

- to Green Monkeys in The Gambia as Revealed by Whole-Genome Sequencing. *Appl Environ Microbiol* 82:5910-7.
335. Turnidge JD, Kotsanas D, Munckhof W, Roberts SA, Bennett CM, Nimmo GR, Coombs GW, Murray RJ, Howden BP, Johnson PDR, Dowling K. 2009. *Staphylococcus aureus* bacteraemia: a major cause of mortality in Australia and New Zealand. *Med J Aust* 191.
 336. Holmes NE, Robinson JO, van Hal SJ, Munckhof WJ, Athan E, Korman TM, Cheng AC, Turnidge JD, Johnson PDR, Howden BP, Vanessa study group obotASfIDCRN. 2018. Morbidity from in-hospital complications is greater than treatment failure in patients with *Staphylococcus aureus* bacteraemia. *BMC Infect Dis* 18:107.
 337. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068-9.
 338. Didelot X, Wilson DJ. 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol* 11:e1004041.
 339. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268-74.
 340. Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10:e1003537.
 341. Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26:1372-3.
 342. Nimmo GR, Bell JM, Collignon PJ, Australian Group for Antimicrobial R. 2003. Fifteen years of surveillance by the Australian Group for Antimicrobial Resistance (AGAR). *Commun Dis Intell Q Rep* 27 Suppl:S47-54.
 343. O'Sullivan CE, Baker MG, Zhang J. 2011. Increasing hospitalizations for serious skin infections in New Zealand children, 1990-2007. *Epidemiol Infect* 139:1794-804.
 344. Stevens DL, Bisno AL, Chambers HF, Everett ED, Dellinger P, Goldstein EJ, Gorbach SL, Hirschmann JV, Kaplan EL, Montoya JG, Wade JC, Infectious Diseases Society of A. 2005. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. *Clin Infect Dis* 41:1373-406.
 345. Williamson DA, Roberts SA, Ritchie SR, Coombs GW, Fraser JD, Heffernan H. 2013. Clinical and molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in New Zealand: rapid emergence of sequence type 5 (ST5)-SCCmec-IV as the dominant community-associated MRSA clone. *PLoS One* 8:e62020.
 346. Kinnevey PM, Shore AC, Brennan GI, Sullivan DJ, Ehricht R, Monecke S, Slickers P, Coleman DC. 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harbouring a novel pseudo staphylococcal cassette chromosome *mec* (SCCmec)-SCC-SCCCRISPR composite element in Irish hospitals. *Antimicrob Agents Chemother* 57:524-31.
 347. Lin YT, Tsai JC, Chen HJ, Hung WC, Hsueh PR, Teng LJ. 2014. A novel staphylococcal cassette chromosomal element, SCCfusC, carrying fusC and speG in fusidic acid-resistant methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 58:1224-7.
 348. Mwangi MW, Wu, S. W., Zhou, Y., Sieradzki, K., de Lancastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E. D., Tomasz, A. 2007. Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc Natl Acad Sci USA* 104:9451-9456.

349. Neoh HM, Cui, L., Yazawa, H., Takeuchi, F., Matsuo, M., Hiramatsu, K. 2008. Mutated response-regulator *graR* is responsible for hetero-VISA to VISA phenotypic conversion of vanomycin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 52:45-53.
350. Nubel U, Dordel, J., Kurt, K., Strommenger, B., Westh, H., Shukla, S. K., Zemlickova, H., Leblois, R., Wirth, T., Jombart, T., Balloux, F., Wittle, E. 2010. A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. *PLoS Pathogens* 6:e1000855.
351. Lindqvist M, Isaksson B, Grub C, Jonassen TO, Hallgren A. 2012. Detection and characterisation of SCCmec remnants in multiresistant methicillin-susceptible *Staphylococcus aureus* causing a clonal outbreak in a Swedish county. *Eur J Clin Microbiol Infect Dis* 31:141-7.
352. Vogel V, Falquet L, Calderon-Copete SP, Basset P, Blanc DS. 2012. Short term evolution of a highly transmissible methicillin-resistant *Staphylococcus aureus* clone (ST228) in a tertiary care hospital. *PLoS One* 7:e38969.
353. Parker D, Narechania A, Sebra R, Deikus G, Larussa S, Ryan C, Smith H, Prince A, Mathema B, Ratner AJ, Kreiswirth B, Planet PJ. 2014. Genome Sequence of Bacterial Interference Strain *Staphylococcus aureus* 502A. *Genome Announc* 2.
354. Giraud C, Hausmann S, Lemeille S, Prados J, Redder P, Linder P. 2015. The C-terminal region of the RNA helicase CshA is required for the interaction with the degradosome and turnover of bulk RNA in the opportunistic pathogen *Staphylococcus aureus*. *RNA Biol* 12:658-74.
355. McCulloch JA, Silveira AC, Lima Moraes Ada C, Perez-Chaparro PJ, Ferreira Silva M, Almeida LM, d'Azevedo PA, Mamizuka EM. 2015. Complete Genome Sequence of *Staphylococcus aureus* FCFHV36, a Methicillin-Resistant Strain Heterogeneously Resistant to Vancomycin. *Genome Announc* 3.
356. Panesso D, Planet PJ, Diaz L, Hugonnet JE, Tran TT, Narechania A, Munita JM, Rincon S, Carvajal LP, Reyes J, Londono A, Smith H, Sebra R, Deikus G, Weinstock GM, Murray BE, Rossi F, Arthur M, Arias CA. 2015. Methicillin-Susceptible, Vancomycin-Resistant *Staphylococcus aureus*, Brazil. *Emerg Infect Dis* 21:1844-8.
357. Azarian T, Maraqa NF, Cook RL, Johnson JA, Bailey C, Wheeler S, Nolan D, Rathore MH, Morris JG, Jr., Salemi M. 2016. Genomic Epidemiology of Methicillin-Resistant *Staphylococcus aureus* in a Neonatal Intensive Care Unit. *PLoS One* 11:e0164397.
358. Harrison EM, Ludden C, Brodrick HJ, Blane B, Brennan G, Morris D, Coll F, Reuter S, Brown NM, Holmes MA, O'Connell B, Parkhill J, Torok ME, Cormican M, Peacock SJ. 2016. Transmission of methicillin-resistant *Staphylococcus aureus* in long-term care facilities and their related healthcare networks. *Genome Med* 8:102.
359. Kpeli G, Darko Otchere I, Lamelas A, Buultjens AL, Bulach D, Baines SL, Seemann T, Giulieri S, Nakobu Z, Aboagye SY, Owusu-Mireku E, Pluschke G, Stinear TP, Yeboah-Manu D. 2016. Possible healthcare-associated transmission as a cause of secondary infection and population structure of *Staphylococcus aureus* isolates from two wound treatment centres in Ghana. *New Microbes New Infect* 13:92-101.
360. Amissah NA, Buultjens AH, Ablordey A, van Dam L, Opoku-Ware A, Baines SL, Bulach D, Tetteh CS, Prah I, van der Werf TS, Friedrich AW, Seemann T, van Dijk JM, Stienstra Y, Stinear TP, Rossen JW.

2017. Methicillin Resistant *Staphylococcus aureus* Transmission in a Ghanaian Burn Unit: The Importance of Active Surveillance in Resource-Limited Settings. *Front Microbiol* 8:1906.
361. Mellmann A, Bletz S, Boking T, Kipp F, Becker K, Schultes A, Prior K, Harmsen D. 2016. Real-Time Genome Sequencing of Resistant Bacteria Provides Precision Infection Control in an Institutional Setting. *J Clin Microbiol* 54:2874-2881.
362. Nubel U, Roumagnac P, Feldkamp M, Song JH, Ko KS, Huang YC, Coombs G, Ip M, Westh H, Skov R, Struelens MJ, Goering RV, Strommenger B, Weller A, Witte W, Achtman M. 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 105:14130-5.
363. Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299-304.
364. Frost LS, Leplae R, Summers AO, Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 3:722-32.
365. MacLean RC, San Millan A. 2015. Microbial Evolution: Towards Resolving the Plasmid Paradox. *Curr Biol* 25:R764-7.
366. Fondi M, Fani R. 2010. The horizontal flow of the plasmid resistome: clues from inter-generic similarity networks. *Environ Microbiol* 12:3228-42.
367. Hulter N, Ilhan J, Wein T, Kadibalban AS, Hammerschmidt K, Dagan T. 2017. An evolutionary perspective on plasmid lifestyle modes. *Curr Opin Microbiol* 38:74-80.
368. Robinson DA, Enright MC. 2003. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 47:3926-34.
369. Gillespie MT, May JW, Skurray RA. 1984. Antibiotic susceptibilities and plasmid profiles of nosocomial methicillin-resistant *Staphylococcus aureus*: a retrospective study. *J Med Microbiol* 17:295-310.
370. Lyon BR, Gillespie MT, Byrne ME, May JW, Skurray RA. 1987. Plasmid-mediated resistance to gentamicin in *Staphylococcus aureus*: the involvement of a transposon. *J Med Microbiol* 23:101-10.
371. Lyon BR, Iuorio JL, May JW, Skurray RA. 1984. Molecular epidemiology of multiresistant *Staphylococcus aureus* in Australian hospitals. *J Med Microbiol* 17:79-89.
372. Wright CL, Byrne ME, Firth N, Skurray RA. 1998. A retrospective molecular analysis of gentamicin resistance in *Staphylococcus aureus* strains from UK hospitals. *J Med Microbiol* 47:173-8.
373. Townsend DE, Ashdown N, Bolton S, Bradley J, Duckworth G, Moorhouse EC, Grubb WB. 1987. The international spread of methicillin-resistant *Staphylococcus aureus*. *J Hosp Infect* 9:60-71.
374. Skurray RA, Rouch DA, Lyon BR, Gillespie MT, Tennent JM, Byrne ME, Messerotti LJ, May JW. 1988. Multiresistant *Staphylococcus aureus*: genetics and evolution of epidemic Australian strains. *J Antimicrob Chemother* 21 Suppl C:19-39.
375. Mahairas GG, Lyon BR, Skurray RA, Pattee PA. 1989. Genetic analysis of *Staphylococcus aureus* with Tn4001. *J Bacteriol* 171:3968-72.
376. Lyon BR, Gillespie MT, Skurray RA. 1987. Detection and characterization of IS256, an insertion sequence in *Staphylococcus aureus*. *J Gen Microbiol* 133:3031-8.
377. Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94.

378. Turnidge JD, Kotsanas D, Munckhof W, Roberts S, Bennett CM, Nimmo GR, Coombs GW, Murray RJ, Howden B, Johnson PD, Dowling K, Australia New Zealand Cooperative on Outcomes in Staphylococcal S. 2009. *Staphylococcus aureus* bacteraemia: a major cause of mortality in Australia and New Zealand. *Med J Aust* 191:368-73.
379. Carver TJ, Rutherford, K. M., Berriman, M., Rajandream, M., Barrell, B. G., Parkhill, J. 2005. ACT: the Artemis comparison tool. *Bioinformatics Appl Notes* 21:3422-3423.
380. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455-77.
381. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691-3.
382. Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070-1.
383. Chen HJ, Lin YT, Hung WC, Tsai JC, Hsueh PR, Teng LJ. 2016. Distribution of Staphylococcal Cassette Chromosome (SCC) *mec* Element Types in Fusidic Acid-Resistant *Staphylococcus epidermidis* and Identification of a Novel SCC7684 Element. *Antimicrob Agents Chemother* 60:5006-9.
384. Butterfield JM, Tsuji BT, Brown J, Ashley ED, Hardy D, Brown K, Forrest A, Lodise TP. 2011. Predictors of agr dysfunction in methicillin-resistant *Staphylococcus aureus* (MRSA) isolates among patients with MRSA bloodstream infections. *Antimicrob Agents Chemother* 55:5433-7.
385. Paulander W, Varming AN, Baek KT, Haaber J, Frees D, Ingmer H. 2012. Antibiotic-mediated selection of quorum-sensing-negative *Staphylococcus aureus*. *mBio* 3:e00459.
386. Amorim ML, Faria NA, Oliveira DC, Vasconcelos C, Cabeda JC, Mendes AC, Calado E, Castro AP, Ramos MH, Amorim JM, de Lencastre H. 2007. Changes in the clonal nature and antibiotic resistance profiles of methicillin-resistant *Staphylococcus aureus* isolates associated with spread of the EMRSA-15 clone in a tertiary care Portuguese hospital. *J Clin Microbiol* 45:2881-8.
387. Hsu LY, Koh TH, Singh K, Kang ML, Kurup A, Tan BH. 2005. Dissemination of multisusceptible methicillin-resistant *Staphylococcus aureus* in Singapore. *J Clin Microbiol* 43:2923-5.
388. Hart J, Christiansen KJ, Lee R, Heath CH, Coombs GW, Robinson JO. 2014. Increased EMRSA-15 health-care worker colonization demonstrated in retrospective review of EMRSA hospital outbreaks. *Antimicrob Resist Infect Control* 3:7.
389. Knight GM, Budd, E. L., Whitney, L., Thornley, A., Al-Ghusein, H., Planche, T., Lindsay, J. A. 2012. Shift in dominant hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) clones over time. *J Antimicrob Chemother* doi:10.1093/jac/dhs245.