Title: Remodelling of pSK1 Family Plasmids and Enhanced Chlorhexidine Tolerance in Methicillin-Resistant *Staphylococcus aureus*

Running Title: Evolutionary Dynamics of pSK1

Authors: Sarah L Baines[1#], Slade O Jensen [2], Neville Firth[3], Anders Gonçalves da Silva[4], Torsten Seemann[5], Glen Carter[4], Deborah A. Williamson[4], Benjamin P Howden[4,5,6#*], Timothy P Stinear[1,5*]

[*] Joint senior authors, # Corresponding authors

Affiliations:

[1]Department of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Victoria, Australia

[2]Microbiology and Infectious Diseases, School of Medicine, Ingham Institute for Applied Medical Research, University of Western Sydney, New South Wales, Australia

[3]School of Life and Environmental Sciences, University of Sydney, New South Wales, Australia

[4]Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Victoria, Australia

[5]Doherty Applied Microbial Genomics, Department of Microbiology & Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Victoria, Australia

[6]Infectious Diseases Department, Austin Health, Victoria, Australia

22 **Word Count:** Abstract (241), Importance (144), Main Text (5,216)

23 **Corresponding authors:**

24 Sarah Baines: bainess@unimelb.edu.au

25 Benjamin Howden: bhowden@unimelb.edu.au

Abstract

Staphylococcus aureus is a significant human pathogen whose evolution and adaptation has been shaped in part by mobile genetic elements (MGEs), facilitating global spread of extensive antimicrobial resistance. However, our understanding of the evolutionary dynamics surrounding MGEs is incomplete, in particular how changes in the structure of multidrug-resistant (MDR) plasmids may influence important staphylococcal phenotypes. Here, we undertook a population- and functional-genomics study of 212 methicillin-resistant *S. aureus* (MRSA) ST239 isolates collected over 32 years to explore the evolution of the pSK1 family of MDR plasmids, illustrating how these plasmids have co-evolved 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, with multiple structural variants arising. Plasmid maintenance and stability was linked to IS*256*- and IS*257*-mediated chromosomal integration and disruption of plasmid replication machinery. Overlaying genomic comparisons with phenotypic susceptibility data for gentamicin and chlorhexidine, it appeared that pSK1 has contributed to enhanced resistance in ST239 MRSA through two mechanisms: (i) acquisition of plasmid-borne resistance mechanisms increasing rates of gentamicin resistance and reduced chlorhexidine susceptibility, and (ii) changes in plasmid configuration linked with further enhancement of chlorhexidine tolerance. While the exact mechanism of enhanced tolerance remains elusive, this research has uncovered a clear evolutionary response of ST239 MRSA to chlorhexidine, one which may contribute to the ongoing persistence and adaptation of this lineage within healthcare institutions.

Importance

The biocide chlorhexidine is fundamental to infection control practices which prevent nosocomial infection and it is highly effective for decolonisation of *S. aureus* from human skin. There have been increasing reports of staphylococcal populations evolving tolerance to chlorhexidine, suggesting that the increasing use and reliance on this biocide may provide a significant selection pressure influencing the evolution of staphylococcal populations. Chlorhexidine tolerance in *S. aureus* is commonly enabled by the acquisition of an efflux pump, however alternative mechanisms influencing tolerance are poorly understood. In this study we demonstrate a previously unrecognised phenomenon, plasmid structural remodelling, by which *S. aureus* may developed enhanced chlorhexidine tolerance. Further, we highlight the importance of undertaking a detailed exploration of the evolutionary dynamics surrounding mobile genetic elements, which contribute immensely to the evolution of microbial species and provide here a framework for how such an analysis can be performed.

Introduction

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 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 and truly pandemic lineage of *S. aureus* is multi-locus sequence type (ST) 239; a multidrug-resistant, healthcare 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, revealing mechanisms which have contributed to its global spread, extensive antimicrobial resistance repertoire, and persistence in healthcare environments [10-15]. In Australia, ST239 has been the dominant HA-MRSA lineage for nearly four decades (national surveillance reports: http://agargroup.org.au/agar-surveys/). Although its prevalence is declining in the regions, having recently been surpassed by the epidemic EMRSA-15 (ST22) clone [16], ST239 is still regularly recovered as a cause of invasive disease in multiple Australian states [17]. We have previously described the long-term evolution of ST239 MRSA in Eastern-Australian hospitals, one of convergent and adaptive evolution of two genetically distinct ST239 clades towards 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 limited exploration of the accessory genome that is primarily composed of MGEs.

Early ST239 MRSA were first recognised in Australia because they displayed resistance to gentamicin [18-20], encoded for by a bifunctional acetyltransferase-phosphotransferase gene

aac(6')-aph(2") commonly carried on pSK1-like plasmids as part of the composite transposon (Tn)*4001* [21, 22]. The focus of multiple studies, pSK1 represents a family of intermediate sized (20-40 kb), theta-replicating staphylococcal plasmids that have been recovered in Australia and the United Kingdom [21, 23-26]. Multiple antibiotic resistance mechanisms are variably encoded for on pSK1-like plasmids. In addition to *aac(6')-aph(2")*, resistance to penicillinase-labile penicillins is encoded for by *blaZ* as part of Tn*552* [24, 25], and trimethoprim resistance is mediated by an insensitive dihydrofolate reductase (dfr), encoded by *dfrA* and carried as part of Tn*4003* which represents a cointegrated remnant of a pSK639-like plasmid [27, 28]. Additionally, harboured by pSK1-like plasmids is a *qacA* gene encoding a quaternary ammonium compound (QAC) multidrug-efflux pump [29], which can mediate tolerance to cationic biocides, most notably chlorhexidine [30]. The presence of a plasmid-borne *qacA* gene, and predicted biocide tolerance, has previously been attributed to an outbreak of ST239 MRSA in the United Kingdom in the early 2000s [31, 32].

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 [33, 34]. It is one of the most widely used antiseptic agents because of it broad spectrum of activity against bacteria, fungi, and enveloped viruses, as well as its good safety record and general tolerability [35-37]. Resistance at in-use concentrations, typically a 0.2 to 4.0% CHX solution in water, has not been reported in *S. aureus*, however the phenomena of enhanced tolerance appears to be increasingly common [38, 39], and is reviewed in [40]. CHX tolerance, defined here as an increase in the minimum inhibitory concentration (MIC) and/or minimum bactericidal concentration (MBC) to a level that remains below in-use concentrations, has been associated with the acquisition or mutation of cationic biocide active efflux pumps, the most commonly reported being

the QAC efflux systems [30, 41, 42]. Most reports of staphylococcal populations evolving CHX tolerance demonstrated either a phenotypic shift in MIC/MBC or detected an increase in the prevalence of genes encoding efflux systems [40, 43]. In *S. aureus* this is mainly associated with *qacA*, often referred to as to *qacA/B* as the encoded peptides differs by only a single amino acid [30, 40, 44]. It is important to note that acquisition of a *qac* gene does not consistently lead to phenotypic tolerance, nor is an increase in CHX tolerance invariably associated with an efflux system [40, 45, 46]. Subsequently, shifts in the prevalence of *qac* genes in staphylococcal populations may not correspond with the development of CHX tolerance. While several studies do combine 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 [38, 43].

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) identify and understand the mechanisms of adaptation with a specific focus on changes in the structural configuration of pSK1-like plasmids, and (ii) explore the phenotypic consequences of these changes in ST239 MRSA, with particular interest in the development of enhanced CHX tolerance.

Results and Discussion

In Australia, the ST239 MRSA population is composed of two genetically distinct clades [15]. The oldest clade is largely restricted to Australia and represents the original lineage circulating in the region. Conversely, the newer clade represents an intercontinental transmission event, introducing the lineage circulating in South East Asia into Australia, a more recent estimate placing this event in the late 1990s (Figure S1). These ST239 clades were originally referred to as Clade 1 and Clade 2 [15], however in this publication we have renamed them as the Australian clade (original) and the Asian-Australian clade (introduced), respectively.

Plasmids related to pSK1 were first identified in ST239 MRSA circulating in Australia in the early 1980s [21, 23, 24]. An *in silico* examination of contemporary isolates revealed that pSK1-like plasmids are still maintained in this population (Figure S1), suggesting the extended co-evolution of pSK1-like plasmids and ST239 MRSA and hence prolonged exposure to the highly selective healthcare environment. To explore the evolutionary dynamics surrounding the pSK1 family we undertook a detailed phenotypic and comparative genomic analysis of 212 temporally (1980 to 2012) and geographically diverse Australian ST239 MRSA. To aid in understanding the changes in plasmid configuration which are described in this publication the structure of the family prototype, pSK1, is illustrated in Figure 1A and described in the Supplementary Results.

Prevalence and Structural Variability of pSK1 Family Plasmids

To identify pSK1-like plasmids, the presence and synteny of plasmid genes was assessed *in silico*. This analysis found that pSK1-like plasmids were solely carried by isolates belonging to the Australian clade, present in 92/124 isolates (74.1%). A comparison of plasmid carriage with year

of isolation illustrated that the prevalence of pSK1-like plasmids had increased over time (Figure 1B), suggesting an evolutionary benefit for their acquisition and maintenance in this population. There was no evidence of plasmid sharing between the Australian and Asian-Australian clades (Figure 2). 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] (Figure 2A).

A closer examination of the 92 isolates in which a pSK1-like plasmid was identified showed considerable variation in the presence of plasmid genes (Figure 2B). When aligned to a model for the ST239 population structure 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. Six broad gene patterns (GP) could be resolved amongst these data (Figure 2B, Supplementary Results). Differences between the GPs correlated with the variable presence of the composite transposons in pSK1. For example, the absence of Tn*4003* in GP3 and GP4 (Figure 2B). Additionally, the loss of three to six syntenic genes in the plasmid backbone, including the replication initiation protein (*repA*) and the plasmid partitioning (*par*) genes (observed in GP2, GP4 and GP6, Figure 2B). This finding suggested that some of the pSK1-like SVs may be chromosomally integrated. Examination of the pSK1 gene presence and synteny in the Asian-Australian clade revealed that 5/6 syntenic genes associated with pSK1 region 4 (including *qacAR*) were highly prevalent (Figure 2B). These isolates were found to carry an alternative *qacA*-containing plasmid, specifically a pTW20_1-like plasmid [31]. Unlike pSK1, the pTW20_1-like gene structure appeared to be largely conserved. A description of this second plasmid population is provided in the Supplementary Results.

Structural Variants of pSK1

To explore the extent of structural remodelling (defined here as changes in plasmid gene content and/or configuration) that has occurred in the pSK1-like plasmid population, long-read sequencing was conducted to establish a reference sequence for the six expected SVs. Additionally, the genome assemblies of all isolates were mined for unique plasmids features, used to determine the prevalence of each SV in the wider ST239 population. This work has been summarised below. For a more detailed explanation refer to the Supplementary Results.

Using this approach eight distinct pSK1-like plasmids were identified, having arisen largely through the activities of IS*256* and/or IS*257*. These changes can be classified into three categories. The first was IS-mediated gain or loss of the composite transposons Tn*4001* and Tn*4003*. This was observed in three SVs, termed SV1, SV3 and SV4 (representing GP1, GP3, and GP4, respectively). SV1 was found to lack the aminoglycoside resistance conferring Tn*4001* and is structurally equivalent to the previously identified pSK7 (Figure 3A) [24, 47]. With only a single copy of the IS*256* target site duplications (TSD) that flank Tn*4001* in pSK1 identified, it was unclear whether SV1 had lost Tn*4001* or had never acquired it. Given when the isolates in which SV1 were recovered (1980 and 1982) and their location close to the tree root in a time-aware phylogenetic model for the Australian clade (Figure S2), SV1 likely represented a progenitor of pSK1 prior to gaining Tn*4001*. Both SV3 and SV4 were found to lack Tn*4003* (Figure 3A,3B). Instead they contained only a single copy of IS*257* with the same TSD that flanked this region in pSK1, in a configuration equivalent to that of pSK14 [24, 47]. Unlike SV1, the phylogenetic location of isolates carrying SV3 and SV4 strongly supported deletion of Tn*4003* (Figure S2), potentially through homologous recombination between the flanking IS*257*s rather than historic absence. This

notion is consistent with the proposed evolution of Tn*4003*, which through IS*257*-mediated transposition from a pSK639-like plasmid was acquired into a pSK1-like precursor to generate the plasmid cointegrate [28].

The second category was IS-mediated chromosomal integration and disruption of the plasmid replication machinery, through IS- and non-IS-mediated deletions/exclusion events. Chromosomal integration was observed in five pSK1-like SVs. Genomic island SV2 (representing GP2) emerged from IS*256*-mediated integration of pSK1 adjacent to the staphylococcal accessory regulator A (*sarA*) gene (Figure 3B). Likewise, SV4 (representing GP4) emerged from IS*256*-mediated integration of SV3 adjacent to a predicted aerobactin biosynthesis gene and in close proximity to an alanine racemase (*alr*) gene (Figure 3B). In both cases, integration was likely preceded by an IS*256* transposition event, resulting in the addition of an extra IS*256* within each plasmid precursor. Interaction between the novel and native IS*256* copies led to the formation of circular intermediates, with each composite Tn encompassing the majority of the plasmid sequence but excluding a segment of the plasmid backbone, resulting in the loss of six and three syntenic CDS in SV2 and SV4, respectively. The three remaining SVs had all arisen from a single IS*257*-mediated chromosomal integration event. Only SV5 and SV6 were represented in the complete genomes; a hypothesised structure has been proposed for the other SV, termed SV5' as it represented the progenitor for both SV5 and SV6 (Figure 3B & 3C). In this case, chromosomal integration of pSK1 had occurred 9.2 kb upstream of a disrupted β-haemolysin gene (disruption resulting from integration of prophage Sa3) and likely involved IS*257*-mediated replicative transposition (for integration) and homologous recombination to account for the partial deletion of Tn*4003*. Genomic island SV5 had an additional 45 bp deletion in *repA* (Figure 3C), and SV6 had a deletion of six syntenic CDS from the plasmid backbone likely resulting from IS*256*-mediated

214 homologous recombination (Figure 3B). All deletion events invariably resulted in the loss of
215 genes, or predicted loss of function of the plasmid replication machinery, which has a known role
216 in stabilising newly integrated elements by removing any interference with chromosomal
217 replication [48].

218 The third category represented two IS-mediated inversion events. The first was an IS*257*-mediated
219 inversion of SV5', giving rise to genomic island SV5 (Figure 3C). This appears to be the result of
220 intramolecular replicative transposition in the inverse orientation and resulted in the inversion of
221 a 42.7 kb region [4]. This has reversed the orientation of all SV5 CDS and split φSa3 (Figure 3C).
222 A second smaller IS*256*-mediated inversion event was identified in SV3. Transposition of IS*256*
223 followed by homologous recombination between the new and a native IS*256* copy (in inverted
224 orientation) has led to a 3.5 kb inversion in the plasmid backbone and reversed the orientation of
225 *repA*, *par* and a hypothetical CDS (Figure 3A). Further, this is the same region deleted in SV4.
226 While the exact consequences of these inversions are unknown the division and partial inversion
227 of φSa3 should prevent excision of the prophage.

228

229 *Convergent Evolution of pSK1-like Variants*

230 Analysis of unique plasmid features in the short-read *de novo* assemblies enabled assignment of
231 all isolates harbouring a pSK1-like plasmid to one of the eight SVs and resolved all inconsistencies
232 between the phylogenetic population model and the originally identified GPs (Supplementary
233 Results). This finding strongly suggested that once a novel pSK1-like plasmid emerged it was
234 maintained and structurally remodelled during clonal expansion, but rarely if ever horizontally
235 transferred between ST239 isolates. To explore potential evolutionary patterns in the plasmid

236 population, the SVs were overlayed on to a Bayesian-inferred time-aware phylogenetic model for
237 the Australian clade (Figure S2).

238 The first pSK1-like plasmids to have appeared in the Australian ST239 clade were SV1 and pSK1,
239 in the late 1970s, consistent with the first identified gentamicin-resistant MRSA in Australia [18-
240 20]. From this model it was estimated that pSK1 emerged in 1979 (95% highest posterior density
241 interval [HPDI]: 1978 - 1980), following the likely acquisition of Tn*4001* into the ancestral SV1
242 plasmid. In the following three decades there was the emergence of SV3 around 1991 (HPDI: 1989
243 – 1994), in which Tn*4003* had been deleted and the replication machinery inverted. There had been
244 at least three independent chromosomal integration events, resulting in the emergence of SV2
245 around 1985 (HPDI: 1983 – 1986), SV5' around 1994 (HPDI: 1991 – 1996), and SV4 around 2000
246 (HPDI: 1996 – 2004). These events likely involved pSK1 as the ancestral plasmid, or SV3 as the
247 immediate ancestor of SV4. Inversion of SV5' in 1999 (HPDI: 1997 – 2002) led to the emergence
248 of SV5. In each integrated variant, the replication machinery had been disrupted through IS*256*-
249 associated gene loss or an internal *repA* deletion. It was originally reported that region 1 was
250 conserved in all pSK1 family plasmids [47], and for the novel extra-chromosomal SVs reported in
251 this study this remains true. However, in the chromosomally integrated SVs only a ~9.3 kb segment
252 of the region is conserved, demonstrating > 95% nucleotide sequence identity with *S. warneri*
253 plasmid pPI-1 (accession AB125341.3) [47, 49]. The genes encompassed within this region encode
254 protein products predicted to be cell-envelope associated, involved in membrane transport and
255 potentially iron acquisition, with one gene encoding a putative Fst-like toxin as part of a Type I
256 toxin-antitoxin system [47]. The near ubiquitous conservation of this region in the pSK1-like
257 plasmid population suggests that these encoded products beneficial and possibly contributing to
258 plasmid maintenance [47].

In addition to the clear stepwise evolution of the pSK1-like population, there also appeared to be evidence of convergent evolution. The three integration events have all occurred independently. Examination of the *de novo* assemblies for all 92 isolates harbouring a pSK1-like plasmid identified a further two independent deletion events amongst the SV5´ and SV5 clade (Figure S2). These findings strongly suggested that the emergence of these phylogenetically distinct but structurally similar SVs is the result of a significant but unknown evolutionary pressure acting on this pSK1-like population. 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 [48]. In the Australian clade, almost all isolates (75/77) that have descended from an ancestral genome with an integrated pSK1-like structure have maintained it (Figure S2). 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 encoded for by genes harboured on pSK1, could also be contributing to the evolutionary pressure promoting plasmid maintenance and driving chromosomal integration.

Evolving Antimicrobial Resistance and Disinfectant Tolerance

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

Trimethoprim. Phenotypic resistance to trimethoprim was detected in all 211 ST239 isolates, with a MIC of > 32 mg/L (Table 1). Genotypically, 75/123 isolates (61.0%) of the Australian clade and

all 88 isolates of the Asian-Australian clade were found to harbour a mutated or acquired *dfr* gene, the *dfrA* harboured by pSK1-like plasmids or *dfrG* respectively (Table 1). A likely mechanism of resistance could not be identified in the remaining 48 isolates from the Australian clade. The lack of phenotypic variation indicated that structural variation in the plasmid population, involving the deletion or alteration of Tn*4003*, had not impacted on trimethoprim resistance in ST239 due to an unidentified genotypic redundancy. Subsequently, trimethoprim is unlikely to be acting as a significant driver of evolution in this population.

Gentamicin. Phenotypic resistance to gentamicin was detected in 185 (87.7%) isolates (Table 1). Three acquired genes encoding aminoglycoside modifying enzymes (AME) were identified [50]: (1) the bifunctional *aac(6')-aph(2")*, found in Tn*4001* and other chromosomal and phage locations [51-53]; (2) an adenyltransferase gene (*aadD*, also known as ANT(4')-Ia) [54, 55]; (3) a phosphotransferase gene (*aph(3')-IIIa*) [56], which has been found co-located with *aac(6')-aph(2")* in φSPβ-like in ST239 *S. aureus* TW20 [31], an isolate closely related to the Asian-Australian clade [15]. The bifunctional AME was common amongst both the Australian and Asian-Australian clades, identified in 98 (79.7%) and 82 (93.2%) isolates, respectively. Conversely, the distribution genes encoding the monofunctional AMEs were more distinct (Table 1). Phenotypically the Asian-Australian clade demonstrated a significant higher average MIC to gentamicin compared to the Australian clade ($p < 0.0001$, Table S1). However, as most isolates harboured multiple AMEs, phenotypic resistance could not be attributed to a single gene. 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 assess the impact of pSK1 evolution on gentamicin MIC, the phenotypic data was investigated for temporal trends. These analyses suggested a trend of increasing gentamicin MIC overtime. However, this was only observed in the Australian clade and not the Asian-Australian clade when modelled separately (Figure S3). This finding is consistent with what has been previously observed in ST239 MRSA with the glycopeptides and daptomycin, hypothesised 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 overtime [15]. These same phenomena have likely contributed to the shift in gentamicin MIC. The Asian-Australian clade had already developed high level gentamicin resistance prior to arriving in Australia, through the acquisition of φSPβ-like 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 MIC (Table S1); the uptake of pSK1 serving as one mechanism by which *aac(6')-aph(2")* could be acquired. However, plasmid evolution, with the emergence of the pSK1-like variants, appeared to have no effect on gentamicin MIC, with no temporal trend having been detected amongst the pSK1-like plasmid containing population (Figure S3, Table S2). Therefore, gentamicin also does not appear to be acting as a significant driver of plasmid structural remodelling.

Chlorhexidine. Phenotypic tolerance to CHX was detected in 150 (71.1%) isolates, defined as an MIC > 2 mg/L [40]. The MIC values of the population ranged from 1 - 6 mg/L, and the MBC values from 2 - 16 mg/L (Table 1). A total of 156 (73.9%) isolates were found to harbour a *qacA* gene and two a *qacC* gene (Table 1). The latter, also known as *qacD*, *smr* or *ebr*, although encoding a biocide active efflux pump is not active against CHX [40]). The *qacA* gene was consistently

associated with isolates that carried either a pSK1-like (Australian clade) or pTW20_1-like plasmid (Asian-Australian clade). The presence of *qacA* was significantly associated with an elevated MIC and MBC to CHX ($p < 0.0001$, Table 2). Although *qacA* was equally prevalent in both the Australian and Asian-Australian clades, 89 (72.4%) and 67 (76.1%) isolates 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, suggesting that genotypic differences between the two plasmid populations could be contributing to the observed phenotypic variation. A comparison of all *qacAR* sequences to reference JKD6008 (Australian clade) identified six SNPs resulting in missense mutations, five in *qacA* and one in *qacR*. Examination of the prevalence, phylogenetic correlation, and phenotypic association of the mutations found that none could explain the difference in tolerance observed between the two clades (Supplementary Results). Using sequence read coverage data, plasmid copy number was also investigated. This analysis revealed that both pSK1-like and pTW20_1-like plasmids were low copy number (average of one and three copies, respectively) and maintained only a single copy of *qacA* per plasmid (Supplementary Results). While this finding was consistent with previous reports and pSK1-like variants commonly being chromosomally integrated [15], it did not explain the observed phenotypic variation and suggested that genotypic difference occurring outside of *qacAR* were responsible.

As with gentamicin, linear models were developed to investigate the CHX phenotypic data for temporal trends, which may indicate a role for pSK1 family evolution in the development of CHX tolerance (Figure S3). These models also suggested a trend of increasing MIC and MBC to CHX overtime. Again, this trend was observed in the Australian but not the Asian-Australian clades when modelled separately. In contrast to the evolutionary phenomena facilitating the development

of reduced antibiotic susceptibility, adaptive evolution in the Australian clade appeared to be the sole contributor to enhanced CHX tolerance. The Asian-Australian clade is less tolerant to CHX, subsequently its introduction into the region, although increasing the prevalence of *qacA*, contributed minimally to the population level shift in phenotype (Table 2). In the Australian clade, the presence of a pSK1-like plasmid was significant associated with increase in CHX MIC and MBC ($p < 0.0001$, Table 2). However, an increase in the prevalence of these plasmids in the population alone could not account for the extent of intra-clade variation observed (Figure 4), and when the plasmid-harbouring population was modelled separately the trend towards enhanced tolerance remained (Figure S3). Comparison of the average MIC and MBC between the pSK1-like variants suggested the more recently emerged SVs (SV2 to SV6) may be associated with increased tolerance (Figure 5, Table S2). When the plasmid-harbouring population were grouped based on structural similarities, chromosomal integration was significantly associated with an increased CHX MIC ($p = 0.0086$, Table S2). Furthermore, a multi-CDS deletion in the plasmid backbone appeared to be associated with a further increase in MIC, although this was statistically non-significant (Table S2). It is unclear why these structural changes were only associated with a shift in MIC, with all groups having demonstrated highly consistent average MBC values (Figure 5), but this might reflect the specific mechanisms mediating enhanced CHX tolerance or the method utilised to estimate the MBC.

Collectively, these findings suggest that convergent evolution of pSK1 is associated with and likely contributing to the development of enhanced CHX tolerance in the Australian clade. Subsequently CHX use, which is extensive in the healthcare environment, being a fundamental component of infection control practices and hand hygiene initiatives in Australia and proven to be effective in reducing rates of invasive staphylococcal disease [57, 58], is a possible driver of pSK1 evolution.

373 Although strongly suggested by these data, this hypothesis will need additional experimentation
374 to confirm this association and demonstrate that exposure to CHX promotes pSK1-like plasmid
375 maintenance through chromosomal integration.

376

377 *Independent Association of CHX Tolerance with pSK1-Like Plasmid Evolution*

378 To examine the strength of the association between plasmid evolution and the development of
379 enhanced CHX tolerance in the ST239 population, in addition to exploring other possible
380 mechanisms that may be contributing to this phenotype, we used three separate statistical-genomic
381 techniques. In contrast to the previous analyses, these models have focused on the CHX phenotype
382 with no prior assumption that plasmid evolution is a contributing factor.

383 Common approaches to associating phenotype with genotype involve techniques such as Genome
384 Wide Association Studies (GWAS) and Discriminant Analysis of Principal Components (DAPC),
385 which look for the presence/absence of mutations and/or genes that are disproportionately
386 correlated with a phenotype of interest. Both approaches were utilised in this work, neither
387 identifying any significant genotypes (outside of the *qacA*-harbouring plasmids) that could be
388 responsible for enhanced CHX tolerance (Supplementary Results).

389 We therefore performed a modified Bayesian phylogeographic analysis, modelling the CHX
390 phenotype rather than geographic location of the ancestral genomes. The hypothesis was that if the
391 development of CHX tolerance was a consequence of plasmid evolution then the ancestral nodes
392 (ANs) from which the SVs were estimated to have emerged should correlate with those in which
393 a shift in MIC is predicted (Figure 5). In this model, a shift in MIC was predicted at the expected
394 ANs for the emergence of the newer SVs. In the case of SV3, SV4 and SV5' the estimated ancestral

MIC had increased to 4 mg/L, from either 2 or 3 mg/L in the preceding ANs. The emergence of SV5 and SV6 also correlated with a shift in MIC from 4 to 6 mg/L. In the older SVs, a shift in MIC was modelled to have occurred a few nodes following the emergence of the SV. This discrepancy is likely due to the presence of multiple isolates not harbouring a pSK1-like plasmid in close proximity to these ANs. In the case of SV1, a shift in MIC from 2 to 3 mg/L was observed in the AN for a sub-clade of isolates. Similarly, a shift in MIC from 2 to 6 mg/L was observed in the AN for a sub-clade of SV2 harbouring isolates. The overall strong correlation between the emergence of the different pSK1-like SVs and estimated shifts in phenotype (5/8 ANs for all SVs, or 7/8 ANs when the SV1 and SV2 sub-clades are considered), in combination with the absence of an identifiable alternative genotypic mechanisms, provides further support for the association between structural remodelling in the pSK1-like plasmid population and the development of enhanced CHX tolerance in ST239 MRSA.

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 co-evolved with the Australian ST239 MRSA lineage for multiple decades. Within the ST239 MRSA population circulating in Australia, the pSK1-like plasmid population is structurally diverse, with eight distinct variants identified having arisen largely through IS*256*- and IS*257*-mediated loss/gain of the composite transposons, chromosomal integration, deletion/exclusion of CDS, and inversions. When assessed by a temporal phylogenetic model, it appeared that the plasmid population had undergone convergent evolution, with the repeated emergence of chromosomally integrated SVs. In investigating potential drivers for plasmid evolution, it was identified that chromosomal integration was strongly associated with the development of enhanced CHX tolerance. While the mechanism mediating enhanced tolerance remains unclear, we speculate that it is linked to altered regulation of the *qacAR* efflux system, potentially resulting from the movement of IS elements and changes in plasmid configuration. These findings support the idea that the widespread and increasing use of CHX is possibly contributing to the evolution of the pSK1 family of plasmids. Although the levels of reduced susceptibility observed in this study remain well below in-use concentrations for this biocide, they do illustrate an evolutionary response in ST239 MRSA, one that may provide an adaptive advantage in healthcare institutions.

Materials and Methods

Bacterial Isolates. This study utilised a temporal (recovered between 1980 and 2012) and geographically diverse collection of 212 Australian ST239 *S. aureus*. Two isolates, including reference *S. aureus* JKD6008, were recovered in New Zealand [59]. All isolates represented cases of clinical infection. To establish global phylogenetic context for the collection, we supplemented this data with the WGS data for a further 319 international ST239 *S. aureus*. Relevant isolate information can be found in Supplementary Dataset.

Whole Genome Sequencing & Sequence Data. The WGS data for 368 of the 531 isolates (73 of the 212 Australian isolates) had been previously published. All sequence data novel to this study has been made publicly available. Seven isolates were subjected to long-read sequencing to enable complete genome assembly. Information about the generation of novel sequence data, relevant WGS information and accession numbers can be found in the Supplementary Materials.

Bioinformatic Analysis. The bioinformatic analyses performed for this study have been explained in detail in the Supplementary Methods and are briefly outlined here. Sequence data was mapped to reference *S. aureus* JKD6008 (GenBank accession CP002120, [60]) or reference plasmid pSK1 (NC_014369, [31]) using Snippy v3.2 (https://github.com/tseemann/snippy). Maximum likelihood phylogenetic trees were generated with IQ-TREE v1.6.1 [61], and maximum clade credibility trees were generated with BEAST v2.4.7 [62]. Trees were visualised 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 was *de novo* assembled using SPAdes v3.11.0 [63], and annotated with Prokka v1.12 [64]. Long-read sequence data was *de novo* assembled using the SMRT Analysis System v2.3.0.140936 (Pacific Biosciences), circularised and reorientated in

449 Geneious v8.1.5 (Biomatters), and polished with Snippy v3.2. Plasmid structural comparisons
450 were conducted using the Artemis Comparison Tool [65]. Ortholog clustering was performed
451 using Roary [66].

452 ***Antimicrobial Susceptibility Testing.*** Phenotypic susceptibility testing to trimethoprim and
453 gentamicin was performed using E-tests (bioMérieux), and interpreted using CLSI guidelines
454 (M100S, 26th Ed). Susceptibility to CHX was performed using a modified broth microdilution
455 method. CHX MICs were read after 24 hours incubation, and all wells were sub-cultured to assess
456 viability. All isolates were tested in biological triplicate and the median values used for statistical
457 analysis. Susceptibility testing procedures are explained in detail in the Supplementary Methods.
458 All statistical analyses were performed in R v3.4.2 (http://www.R-project.org/), with significance
459 determined as a p value ≤ 0.05.

460 **Acknowledgements.**

461 This work was funded by the National Health and Medical Research Council, Australia
462 (Fellowships GNT1105905 to BPH; Project GNT457454 to NF and SOJ). SLB was supported by
463 an Australian Government Research Training Program Scholarship, and a Victorian Fellowship
464 provided by the Victorian State Government.

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

466

467 **Figure Legends.**

468 **Figure 1. Structure and prevalence of pSK1 plasmids in ST239 MRSA.** (A) pSK1 sequence
469 and annotations are that previously published, GenBank accession GU565967.1 [47]. Predicted
470 CDS have been coloured based on defined regions. Insertion sequences (IS) are coloured grey
471 (IS*256*) and black (IS*257*), with target site duplications (TSD) illustrated: arrows indicate
472 upstream/downstream sequences, orientation, and are coloured to represent unique sequences
473 (refer to key). (B) Graph illustrates the increasing prevalence of pSK1 family plasmids in the
474 Australian ST239 clade overtime. The cumulative total of the sampled population is indicated by
475 the solid line and the proportion in which a pSK1-like plasmid was identified by the broken line.

476 **Figure 2. pSK1 plasmid gene presence and synteny.** (A) Maximum likelihood phylogenetic tree
477 inferred from 3,883 core genome SNPs illustrates the population structure of ST239 *S. aureus* in
478 Australia. Tips are coloured based on location (refer to key). Branches with < 70% bootstrap
479 support are coloured red. (B) Coloured blocks represent the identification of a pSK1 gene, using a
480 95% amino acid homology threshold (excluding IS elements). Box length is reflective of gene

length and ordered based on pSK1 (Figure 1A). Boxes are linked if CDS were found to be syntenic. Coloured boxes reflect one of six defined gene patterns (GP), with location of GP label indicating the isolates selected for long-read sequencing.

Figure 3. pSK1-like structural variants. Illustrated is a schematic of the development of the pSK1-like structural variants. Predicted CDS are coloured based on defined pSK1 regions (Figure 1A), IS*256* are grey, IS*257* are black, and chromosomal CDS are lilac. Arrows denote upstream/downstream target site duplications (TSD), and direction denotes IS orientation. Circles represent a single copy of a TSD not adjacent to an IS element. Arrows and circles are coloured to reflect unique sequences (refer to key), and a star has been used to indicate a TSD present in the reverse complement to what was expected. In the structural comparisons, connected regions share ≥ 98% nucleotide sequence identity, coloured pink or blue to indicate matching or reverse orientation, respectively. (A) Illustrates the emergence of pSK1-like variants through IS-mediated loss/gain of the composite transposons: (i) Tn*4001* acquired in SV1 to produce pSK1, and (ii) Tn*4003* deleted from pSK1 to produce SV3 (with an additional IS*256*-mediated inversion). (B) Illustrates three IS-mediated chromosomal integration events: (i) integration of pSK1 adjacent to *sarA* (SV2), integration of SV3 near *alr* (SV4), and integration of pSK1 near φSa3 (SV6). All three SVs have undergone IS*256*-mediated exclusion/deletion of CDS encoding the plasmid replication machinery. (C) Illustrates the large IS*257*-mediated inversion event in the hypothesised SV5' that gave rise to SV5 and resulted in the fragmentation of φSa3.

Figure 4. Phenotypic variation in chlorhexidine tolerance. Graphs illustrate the distribution of chlorhexidine MIC (top panel) and MBC (bottom panel) values in the Australian clade. Boxplot features represent the population median (central black line), upper and lower quartiles (box), and range (bars) excluding outliers (circles). Boxplots representing SVs are coloured to reflect a

plasmid structural feature: extra-chromosomal plasmid (teal), and chromosomally integrated with either an internal *repA* deletion (light blue) or a multi-CDS region 1 deletion (dark blue).

Figure 5. Bayesian phylogenetic model associating chlorhexidine tolerance with pSK1-like plasmid evolution. Illustrated is a maximum clade credibility tree inferred from the whole genome alignment of the Australian clade (n = 124). Isolates identified as harbouring a pSK1-like plasmid are indicated by a circle located adjacent to the tree and coloured based on the SV identified. The ancestral nodes in which each SV is estimated to have emerged are indicated by a number (refer to key), those coloured black represent an extra-chromosomal plasmid and those coloured white represent a genomic island. The estimated CHX MIC for all ancestral nodes is indicated by a circle, coloured based on the MIC value and sized according to the posterior probability for the estimate (refer to key). Blue bars represent the 95% highest posterior density interval for the node heights. The aligned heatmap illustrates the phenotypic MIC values attained for each isolate.

Table 1. Population Distributions of Antimicrobial and Biocide Resistance Genes and Phenotypic Resistance Profiles.

	All ST239 (n = 211)	Asian-Australian Clade (n = 88)	Australian Clade (n = 123)	SK1 Plasmid (n = 91)
Phenotypic antimicrobial resistance - median (range) mg/L				
Trimethoprim (MIC)[a]	> 32 (-)	-	-	-
Gentamicin (MIC)[a]	32 (0.023 – 256)	256 (0.38 – 256)	16 (0.023 – 256)	24 (0.032 – 256)
Chlorhexidine (MIC)[b]	3 (1 – 6)	3 (1.5 – 4)	4 (1 – 6)	4 (1 – 6)
Chlorhexidine (MBC)[b]	6 (2 – 16)	6 (3 – 12)	8 (2 – 16)	8 (2 – 16)
Prevalence of acquired resistance genes[c] (n)				
***dfr* genes**	*dfrA* (76)	-	*dfrA* (76)	*dfrA* (76)
	dfrG (88)	*dfrG* (88)	-	-
AME genes	*aac6'-aph2"* (180)	*aac6'-aph2"* (82)	*aac6'-aph2"* (98)	*aac6'-aph2"* (85)
	aadD (72)	-	*aadD* (72)	*aadD* (67)
	aph(3')-III (95)	*aph(3')-III* (87)	*aph3'-III* (8)	*aph3'-III* (5)
***qac* genes**	*qacA* (156)	*qacA* (67)	*qacA* (89)	*qacA* (89)
	qacC (2)	-	*qacC* (2)	-

Abbreviations: MIC, minimum inhibitory concentration; MBC, minimum bactericidal concentration

[a] Phenotypic susceptibility against trimethoprim and gentamicin was determined by E-test.

[b] Phenotypic susceptibility against chlorhexidine was determined by broth microdilution.

[c] Identification of resistance genes was determined by local alignment, minimum alignment of 70% gene length and > 95% nucleotide homology required to call a match.

Table 2. Investigation of Chlorhexidine Tolerance in Australian ST239 MRSA

Isolate Populations (n)	Median (mg/L) MIC; MBC		Isolate Populations (n)	Median (mg/L) MIC; MBC	*P* value MIC; MBC
All ST239 (211)					
***qacA* Negative (55)**	0.90; 2.30	vs	***qacA* Positive (156)**	3.70; 8.00	< 0.0001; < 0.0001
Australian Clade (123)	3.40; 7.30	vs	**Asian-Australian Clade (88)**	2.30; 6.70	< 0.0001; NS
SK1 plasmid (91)	4.20; 8.60	vs	**pTW20_1-like (67)**	3.00; 7.20	< 0.0001; 0.0021
Asian-Australian Clade (88)					
***qacA* Negative (21)**	1.90; 5.20	vs	***qacA* Positive (67)**	3.00; 7.20	< 0.0001; < 0.0001
Australian Clade (123)					
***qacA* Negative (34)**	1.80; 4.60	vs	***qacA* Positive (89)**	4.30; 8.60	< 0.0001; < 0.0001
Extra-chromosomal SVs (17)	3.20; 7.80	vs	**Integrated SV (74)**	4.50; 8.74	0.0086; NS
Integrated SV + *repA* deletion (25)[b]	3.70; 8.50	vs	**Integrated SV + R1 deletion (49)[c]**	4.70; 8.60	NS; NS

Abbreviations: MIC, minimum inhibitory concentration; MBC, minimum bactericidal concentration; R1, pSK1 region 1; SV, structural variant.

[a] Population includes SV5' and SV5; [b] Population includes SV2, SV4, and SV6.

Supplementary Materials

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 Methods. This file contains additional information about materials and methods used in this study.

Supplementary Tables. This file contains additional tables which provide information about the distribution of gentamicin and chlorhexidine susceptibility in the ST239 MRSA population.

Supplementary Results. This file contains additional analyses and results relevant to this study.

Supplementary Figure S1. Global population structure of ST239 *S. aureus.* Illustrated is a maximum clade credibility tree inferred from the whole genome alignment of the 531 international ST239 isolates. Tips are coloured based on location (refer to keys). Nodes with > 95% posterior support are indicated by a red dot. The Australian and two Asian-Australian clades (major and minor) and estimates for the most recent common ancestor (MRCA) are indicated, displayed as "median year (95% highest posterior density range)".

Supplementary Figure S2. Phylogenetic model for the emergence of pSK1 plasmid variants. Illustrated is a maximum clade credibility tree inferred from the whole genome alignment of the Australian clade (n = 124). Isolates identified as harbouring a pSK1-like plasmid are indicated by a circle located at the branch tip and coloured based on the SV identified (refer to key). The most recent common ancestor (MRCA) for each SV is indicated by the larger node circle, with black numbers indicating an extrachromosomal SV and white numbers a chromosomally integrated SV. Temporal estimates for these nodes have been provided, displayed as "SV – median year (95%

highest posterior density range)". The two isolates with pSK1 region 1 (R1) deletions in the SV5' and SV5 clade are indicated by a star. Grey arrows illustrate the likely order of structural remodelling that has occurred during the evolution of the pSK1-like plasmid population. Branches with < 95% posterior support are coloured red. The 95% highest posterior density for node heights is represented by the blue bars.

Supplementary Figure 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 = 96).

References

1. **Ochman H, Lawrence JG, Groisman EA**. Lateral gene transfer and the nature of bacterial innovation. *Nature* 2000;405(6784):299-304.
2. **Frost LS, Leplae R, Summers AO, Toussaint A**. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 2005;3(9):722-732.
3. **Soucy SM, Huang J, Gogarten JP**. Horizontal gene transfer: building the web of life. *Nat Rev Genet* 2015;16(8):472-482.
4. **Partridge SR, Kwong SM, Firth N, Jensen SO**. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* 2018;31(4).
5. **van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL et al.** Predictors of mortality in *Staphylococcus aureus* Bacteremia. *Clin Microbiol Rev* 2012;25(2):362-386.
6. **Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr.** *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* 2015;28(3):603-661.
7. **Pavillard R, Harvey K, Douglas D, Hewstone A, Andrew J et al.** Epidemic of hospital-aqcuired infection due to methicillin-resistant *Staphylococcus aureus* in major Victorian hospitals. *Med J Aust* 1982;29(1):451-454.
8. **Cookson BD, Philips I**. Epidemic methicillin-resistant *Staphylococcus aureus*. *J Antimicrob Chemother* 1988;21 (Suppl.C):57-65.
9. **Dubin DT, Chikramane SG, Inglis B, Matthews PR, Stewart PR**. Physical mapping of the *mec* region of an Australian methicillin-resistant *Staphylococcus aureus* lineage and a closely related American strain. *J Gen Microbiol* 1992;138:169-180.
10. **Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK et al.** Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 2010;327:469-474.
11. **Castillo-Ramirez S, Corander J, Marttinen P, Aldeljawi M, Hanage WP et al.** Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biol* 2012;13:R126.
12. **Gray RR, Tatem AJ, Johnson JA, Alekseyenko AV, Pybus OG et al.** Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a bayesian framework. *Mol Biol Evol* 2011;28(5):1593-1603.
13. **Hsu LY, Harris SR, Chlebowicz MA, Lindsay JA, Koh TH et al.** Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol* 2015;16:81.
14. **Tong SY, Holden MT, Nickerson EK, Cooper BS, Koser CU et al.** Genome sequencing defines phylogeny and spread of methicillin-resistant *Staphylococcus aureus* in a high transmission setting. *Genome Res* 2015;25(1):111-118.
15. **Baines SL, Holt KE, Schultz MB, Seemann T, Howden BO et al.** Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*. *MBio* 2015;6(2):e00080.
16. **Coombs GW, Nimmo GR, Pearson JC, Collignon PJ, Bell JM et al.** Australian Group on Antimicrobial Resistance Hospital-onset *Staphylococcus aureus* Surveillance Programme annual report, 2011. *Commun Dis Intell Q Rep* 2013;37(3):E210-218.
17. **Coombs GW, Daley DA, Thin Lee Y, Pearson JC, Robinson JO et al.** Australian Group on Antimicrobial Resistance Australian Staphylococcus aureus Sepsis Outcome Programme annual report, 2014. *Commun Dis Intell Q Rep* 2016;40(2):E244-254.
18. **Graham DR, King K, Brady LM, Karkness J**. Gentamicin-Resistant *Staphylococci*. *Lancet* 1981;318(8248):698-699.

19. **Gedney J, Lacey RW**. Properties of methicillin-resistant staphylococci now endemic in Australia. *Med J Aust* 1982;1(11):448-450.
20. **Gilbert GL, Asche V, Hewstone AS, Mathiesen JL**. Methicillin-resistant *Staphylococcus aureus* in neonatal nurseries. Two years' experience in special-care nurseries in Melbourne. *Med J Aust* 1982;1(11):455-459.
21. **Gillespie MT, May JW, Skurray RA**. Antibiotic susceptibilities and plasmid profiles of nosocomial methicillin-resistant *Staphylococcus aureus*: a retrospective study. *J Med Microbiol* 1984;17(3):295-310.
22. **Lyon BR, May JW, Skurray RA**. Tn*4001*: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Mol Gen Genet* 1984;193(3):554-556.
23. **Lyon BR, Iuorio JL, May JW, Skurray RA**. Molecular epidemiology of multiresistant *Staphylococcus aureus* in Australian hospitals. *J Med Microbiol* 1984;17(1):79-89.
24. **Skurray RA, Rouch DA, Lyon BR, Gillespie MT, Tennent JM et al.** Multiresistant *Staphylococcus aureus*: genetics and evolution of epidemic Australian strains. *J Antimicrob Chemother* 1988;21 Suppl C:19-39.
25. **Wright C, Byrne M, Firth N, Skurray R**. A reterospective molecular analysis of gentamicin resistant in *Staphylococcus aureus* strains from UK hospitals. *J Med Microbiol* 1998;47:173-178.
26. **Townsend DE, Ashdown N, Bolton S, Bradley J, Duckworth G et al.** The international spread of methicillin-resistant *Staphylococcus aureus*. *J Hosp Infect* 1987;9(1):60-71.
27. **Rouch DA, Messerotti LJ, Loo LS, Jackson CA, Skurray RA**. Trimethoprim resistance transposon Tn*4003* from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS*257*. *Mol Microbiol* 1989;3(2):161-175.
28. **Firth N, Skurray RA**. Mobile elements in the evolution and spread of multiple-drug resistance in staphylococci. *Drug Resist Updat* 1998;1(1):49-58.
29. **Tennent JM, Lyon BR, Midgley M, Jones IG, Purewal AS et al.** Physical and biochemical characterization of the *qacA* gene encoding antiseptic and disinfectant resistance in *Staphylococcus aureus*. *J Gen Microbiol* 1989;135(1):1-10.
30. **Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA**. 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* 1996;93(8):3630-3635.
31. **Holden MTG, Lindsay JA, Corton C, Quail MA, Cockfield JD et al.** 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* 2010;192(3):888-892.
32. **Batra R, Cooper BS, Whiteley C, Patel AK, Wyncoll D et al.** 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* 2010;50(2):210-217.
33. **Septimus EJ, Schweizer ML**. Decolonization in Prevention of Health Care-Associated Infections. *Clin Microbiol Rev* 2016;29(2):201-222.
34. **Davies GE, Francis J, Martin AR, Rose FL, Swain G**. 1:6-Di-4'-chlorophenyldiguanidohexane (hibitane); laboratory investigation of a new antibacterial agent of high potency. *Br J Pharmacol Chemother* 1954;9(2):192-196.
35. **Hugo WB, Longworth AR**. Some Aspects of the Mode of Action of Chlorhexidine. *J Pharm Pharmacol* 1964;16:655-662.
36. **Russell AD**. Chlorhexidine: antibacterial action and bacterial resistance. *Infection* 1986;14(5):212-215.
37. **Kampf G, Kramer A**. Epidemiologic background of hand hygiene and evaluation of the most important agents for scrubs and rubs. *Clin Microbiol Rev* 2004;17(4):863-893, table of contents.
38. **Hardy K, Sunnucks K, Gil H, Shabir S, Trampari E et al.** Increased Usage of Antiseptics Is Associated with Reduced Susceptibility in Clinical Isolates of *Staphylococcus aureus*. *MBio* 2018;9(3).

653 39. **Hayden MK, Lolans K, Haffenreffer K, Avery TR, Kleinman K et al.** Chlorhexidine and Mupirocin
654 Susceptibility of Methicillin-Resistant *Staphylococcus aureus* Isolates in the REDUCE-MRSA Trial. *J Clin*
655 *Microbiol* 2016;54(11):2735-2742.
656 40. **Horner C, Mawer D, Wilcox M**. Reduced susceptibility to chlorhexidine in staphylococci: is it
657 increasing and does it matter? *J Antimicrob Chemother* 2012;67(11):2547-2559.
658 41. **Russell AD**. Do biocides select for antibiotic resistance? *J Pharm Pharmacol* 2000;52(2):227-233.
659 42. **Maillard JY**. Bacterial resistance to biocides in the healthcare environment: should it be of
660 genuine concern? *J Hosp Infect* 2007;65 Suppl 2:60-72.
661 43. **Williamson DA, Carter GP, Howden BP**. Current and Emerging Topical Antibacterials and
662 Antiseptics: Agents, Action, and Resistance Patterns. *Clin Microbiol Rev* 2017;30(3):827-860.
663 44. **Littlejohn TG, Paulsen IT, Gillespie MT, Tennent JM, Midgley M et al.** Substrate specificity and
664 energetics of antiseptic and disinfectant resistance in *Staphylococcus aureus*. *FEMS Microbiol Lett*
665 1992;74(2-3):259-265.
666 45. **Harkins CP, McAleer MA, Bennett D, McHugh M, Fleury OM et al.** The widespread use of
667 topical antimicrobials enriches for resistance in Staphylococcus aureus isolated from patients with atopic
668 dermatitis. *Br J Dermatol* 2018;179(4):951-958.
669 46. **Skovgaard S, Larsen MH, Nielsen LN, Skov RL, Wong C et al.** Recently introduced qacA/B genes
670 in Staphylococcus epidermidis do not increase chlorhexidine MIC/MBC. *J Antimicrob Chemother*
671 2013;68(10):2226-2233.
672 47. **Jensen SO, Apisiridej S, Kwong SM, Yang YH, Skurray RA et al.** Analysis of the prototypical
673 *Staphylococcus aureus* multiresistance plasmid pSK1. *Plasmid* 2010;64:135-142.
674 48. **Hulter N, Ilhan J, Wein T, Kadibalban AS, Hammerschmidt K et al.** An evolutionary perspective
675 on plasmid lifestyle modes. *Curr Opin Microbiol* 2017;38:74-80.
676 49. **Aso Y, Sashihara T, Nagao J, Kanemasa Y, Koga H et al.** Characterization of a gene cluster of
677 *Staphylococcus warneri* ISK-1 encoding the biosynthesis of and immunity to the lantibiotic, nukacin ISK-
678 1. *Biosci Biotechnol Biochem* 2004;68(8):1663-1671.
679 50. **Ramirez MS, Tolmasky ME**. Aminoglycoside modifying enzymes. *Drug Resist Updat*
680 2010;13(6):151-171.
681 51. **Gillespie MT, Lyon BR, Messerotti LJ, Skurray RA**. Chromosome- and plasmid-mediated
682 gentamicin resistance in *Staphylococcus aureus* encoded by Tn*4001*. *J Med Microbiol* 1987;24(2):139-
683 144.
684 52. **Mahairas GG, Lyon BR, Skurray RA, Pattee PA**. Genetic analysis of *Staphylococcus aureus* with
685 Tn*4001*. *J Bacteriol* 1989;171(7):3968-3972.
686 53. **Rouch DA, Byrne, M. E., Kong, Y. C., Skurray, R. A.** The *aacA-aphD* gentamicin and kanamycin
687 resistance determinant of Tn*4001* from *Staphylococcus aureus*: expression and nucleotide sequence
688 analysis. *J Gen Microbiol* 1987;133:3039-3052.
689 54. **McKenzie T, Hoshino T, Tanaka T, Sueoka N**. The nucleotide sequence of pUB110: some salient
690 features in relation to replication and its regulation. *Plasmid* 1986;15(2):93-103.
691 55. **Santanam P, Kayser FH**. Purification and characterization of an aminoglycoside inactivating
692 enzyme from *Staphylococcus epidermidis* FK109 that nucleotidylates the 4'- and 4''-hydroxyl groups of
693 the aminoglycoside antibiotics. *J Antibiot (Tokyo)* 1978;31(4):343-351.
694 56. **Trieu-Cuot P, Courvalin P**. Nucleotide sequence of the *Streptococcus faecalis* plasmid gene
695 encoding the 3'5''-aminoglycoside phosphotransferase type III. *Gene* 1983;23(3):331-341.
696 57. **Grayson ML, Russo PL, Cruickshank M, Bear JL, Gee CA et al.** Outcomes from the first 2 years of
697 the Australian National Hand Hygiene Initiative. *Med J Aust* 2011;195(10):615-619.
698 58. **Johnson PD, Martin R, Burrell LJ, Grabsch EA, Kirsa SW et al.** Efficacy of an
699 alcohol/chlorhexidine hand hygiene program in a hospital with high rates of nosocomial methicillin-
700 resistant Staphylococcus aureus (MRSA) infection. *Med J Aust* 2005;183(10):509-514.

59. **Howden BP, Johnson PDR, Ward PB, Stinear TP, Davies JK**. Isolates with low-level vancomycin resistance associated with persistent methicillin-resistant *Staphylococcus aureus* bacteremia. *Antimicrob Agents Chemother* 2006;50(9):3039-3047.

60. **Howden BP, Seemann T, Harrison PF, McEvoy CR, Stanton J-AL et al.** Complete genome sequence of *Staphylococcus aureus* strain JKD6008, an ST239 clones of methicillin-resistant *Staphylococcus aureus* with intermediate level vancomycin resistance. *J Bacteriol* 2010;192(21):5848-5849.

61. **Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ**. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32(1):268-274.

62. **Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH et al.** BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 2014;10(4):e1003537.

63. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19(5):455-477.

64. **Seemann T**. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30(14):2068-2069.

65. **Carver TJ, Rutherford, K. M., Berriman, M., Rajandream, M., Barrell, B. G., Parkhill, J.** ACT: the Artemis comparison tool. *Bioinformatics Appl Notes* 2005;21(16):3422-3423.

66. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S et al.** Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31(22):3691-3693.

Figure 1

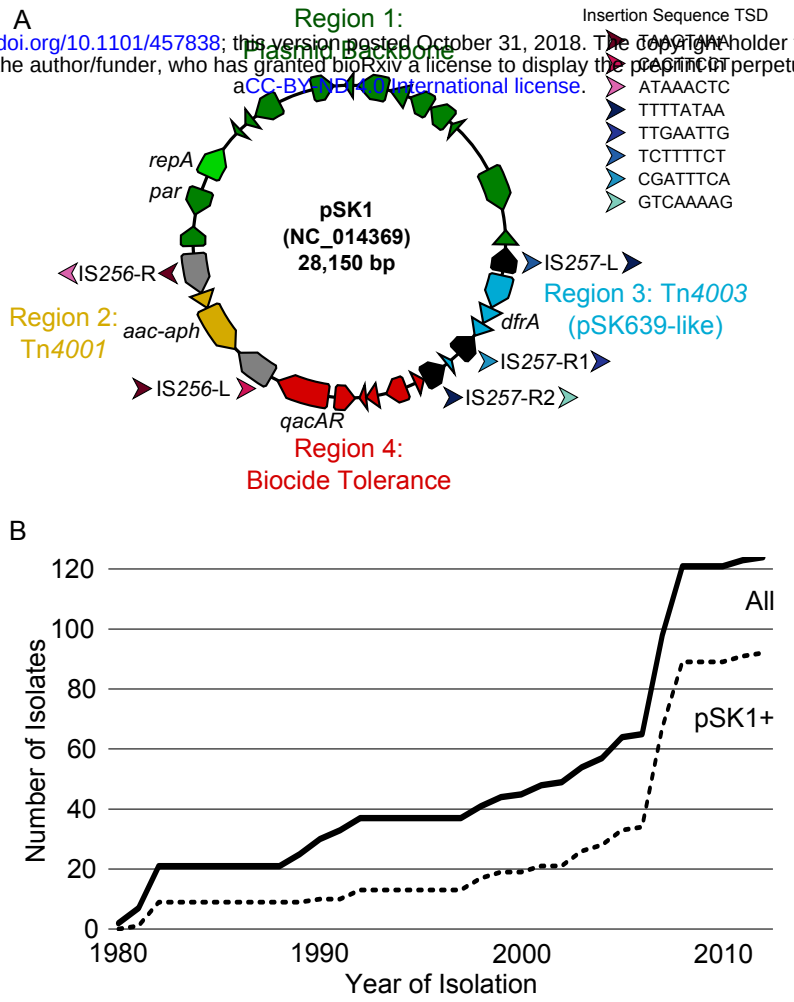

Figure 1. Structure and prevalence of pSK1 plasmids in ST239 MRSA. (A) pSK1 sequence and annotations are that previously published, GenBank accession GU565967.1 [47]. Predicted CDS have been coloured based on defined regions. Insertion sequences (IS) are coloured grey (IS256) and black (IS257), with target site duplications (TSD) illustrated: arrows indicate upstream / downstream sequences, orientation, and are coloured to represent unique sequences (refer to key). (B) Graph illustrates the increasing prevalence of pSK1 family plasmids in the Australian ST239 clade overtime. The cumulative total of the sampled population is indicated by the solid line and the proportion in which a pSK1-like ike plasmid was identified by the broken line.

Figure 2

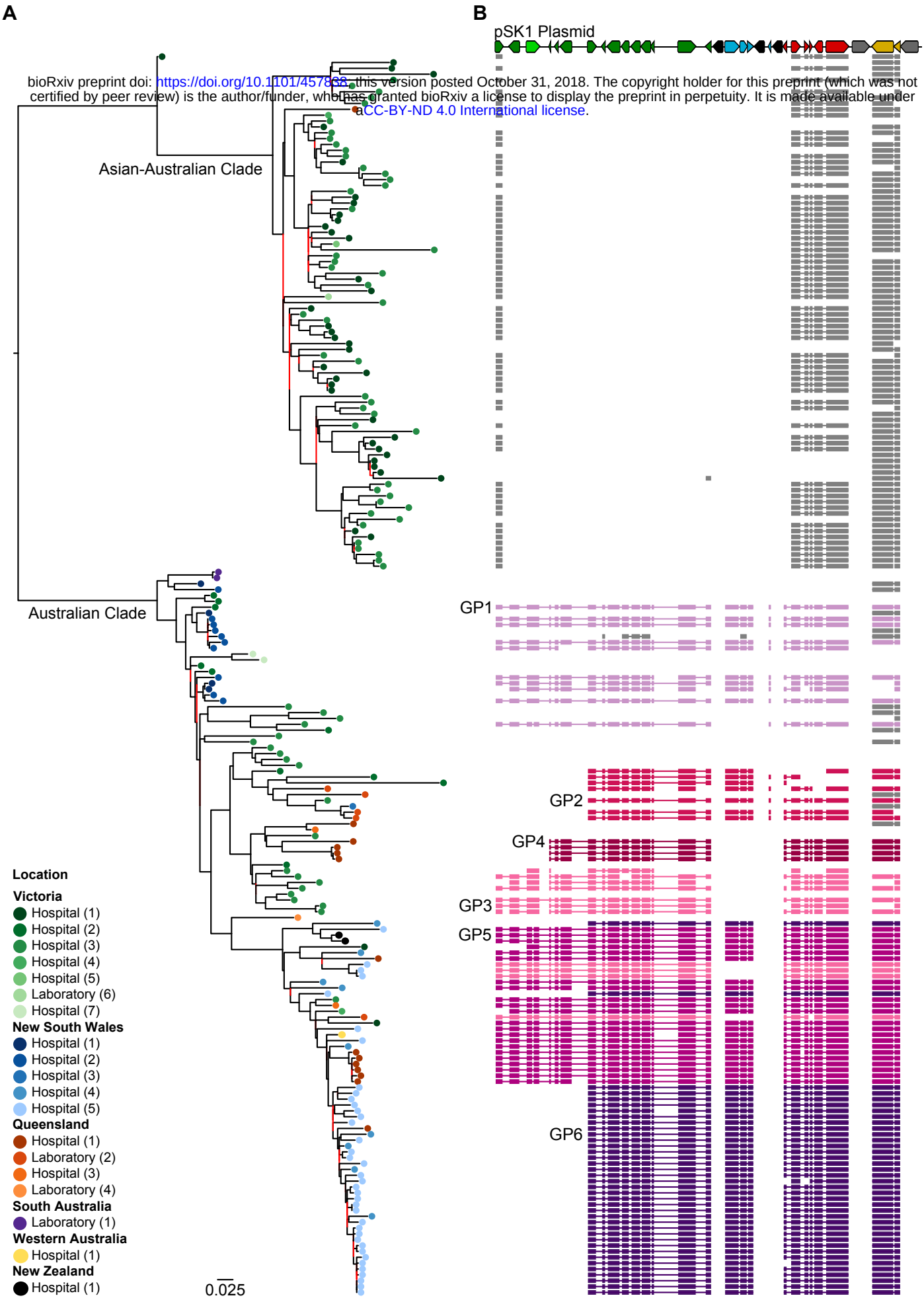

Figure 2. pSK1 SK1 plasmid gene presence and synteny. (A) Maximum likelihood phylogenetic tree inferred from 3,883 core genome SNPs illustrates the population structure of ST239 *S. aureus* in Australia. Tips are coloured based on location (refer to key). Branches with < 70% bootstrap support are coloured red. (B) Coloured blocks represent the identification of a pSK1 gene, using a 95% amino acid homology threshold (excluding IS elements). Box length is reflective of gene length and ordered based on pSK1 (Figure 1A). Boxes are linked if CDS were found to be syntenic. Coloured boxes reflect one of six defined gene patterns (GP), with location of GP label indicating the isolates selected for long-read sequencing.

Figure 3

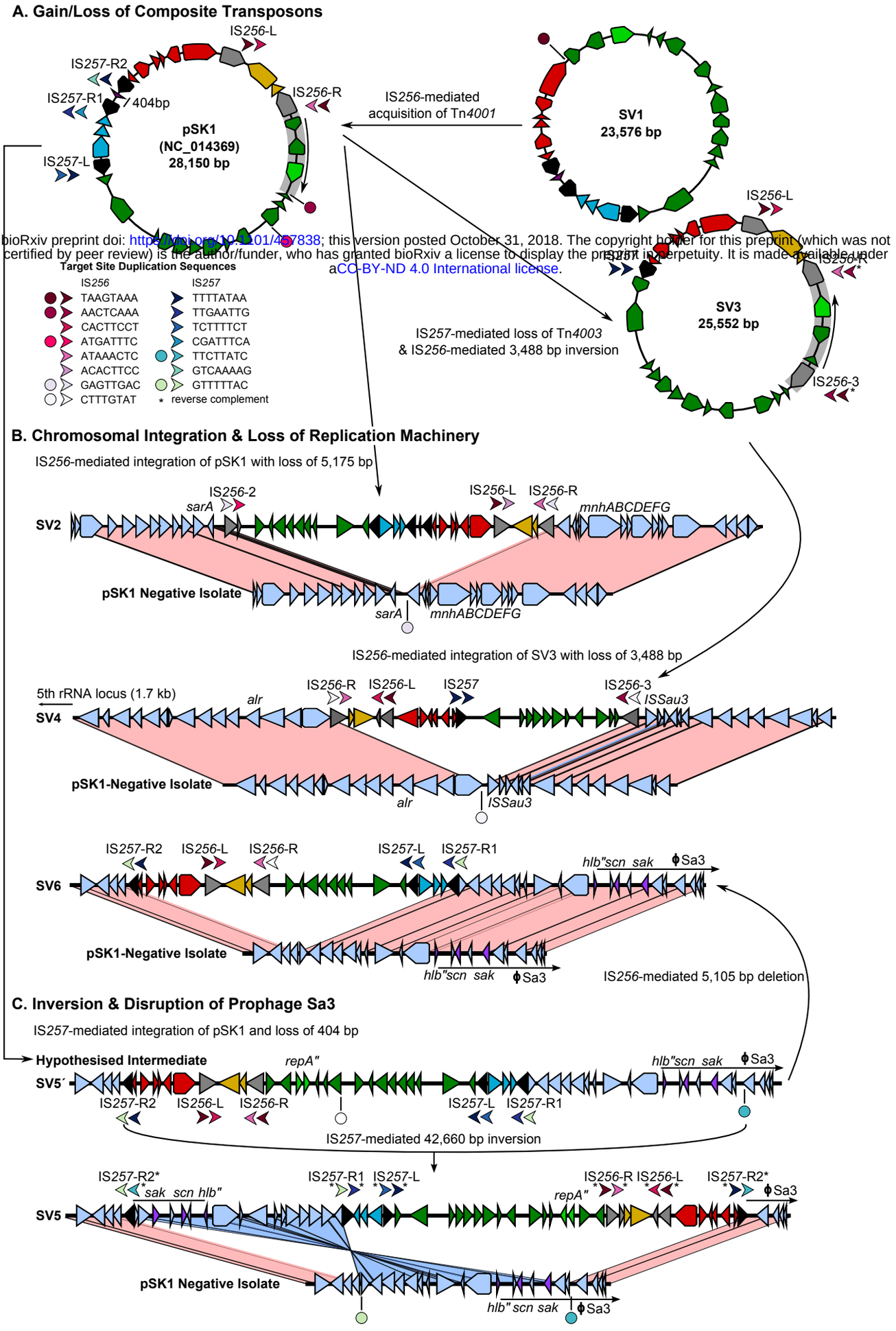

Figure 3. pSK1-like structural variants. Illustrated is a schematic for the development of the pSK1-like plasmid variants. Predicted CDS are coloured based on defined pSK1 regions (Figure 1A), IS*256* are grey, IS*257* are black, and chromosomal CDS are lilac. Arrows denote upstream/downstream target site duplications (TSD) and direction denotes IS orientation. Circles represent a single copy of a TSD not adjacent to an IS element. Arrows and circles are coloured to reflect unique sequences (refer to key) and a star has been used to indicate a TSD present in the reverse complement to what was expected. In the structural comparisons, connected regions share ≥ 98% nucleotide sequence identity, coloured pink or blue to indicate matching or reverse o rientation, respectively. (A) Illustrates the emergence of pSK1-like variants through IS-mediated loss/gain of the composite transposons: (i) Tn*4001* acquired in SV1 to produce pSK1, and (ii) Tn*4003* deleted from pSK1 to produce SV3 (with an additional IS*256*-mediated inversion). (B) Illustrates three IS-mediated chromosomal integration events: (i) integration of pSK1 adjacent to *sarA* (SV2), integration of SV3 near *alr* (SV4), and integration of pSK1 near φSa3 (SV6). All three SVs have undergone IS*256*-mediated exclusion/deletion of CDS encoding the plasmid replication machinery. (C) Illustrates the large IS*257*-mediated inversion event in the hypothesised SV5' that gave rise to SV5 and resulted in the fragmentation of φSa3.

Figure 4

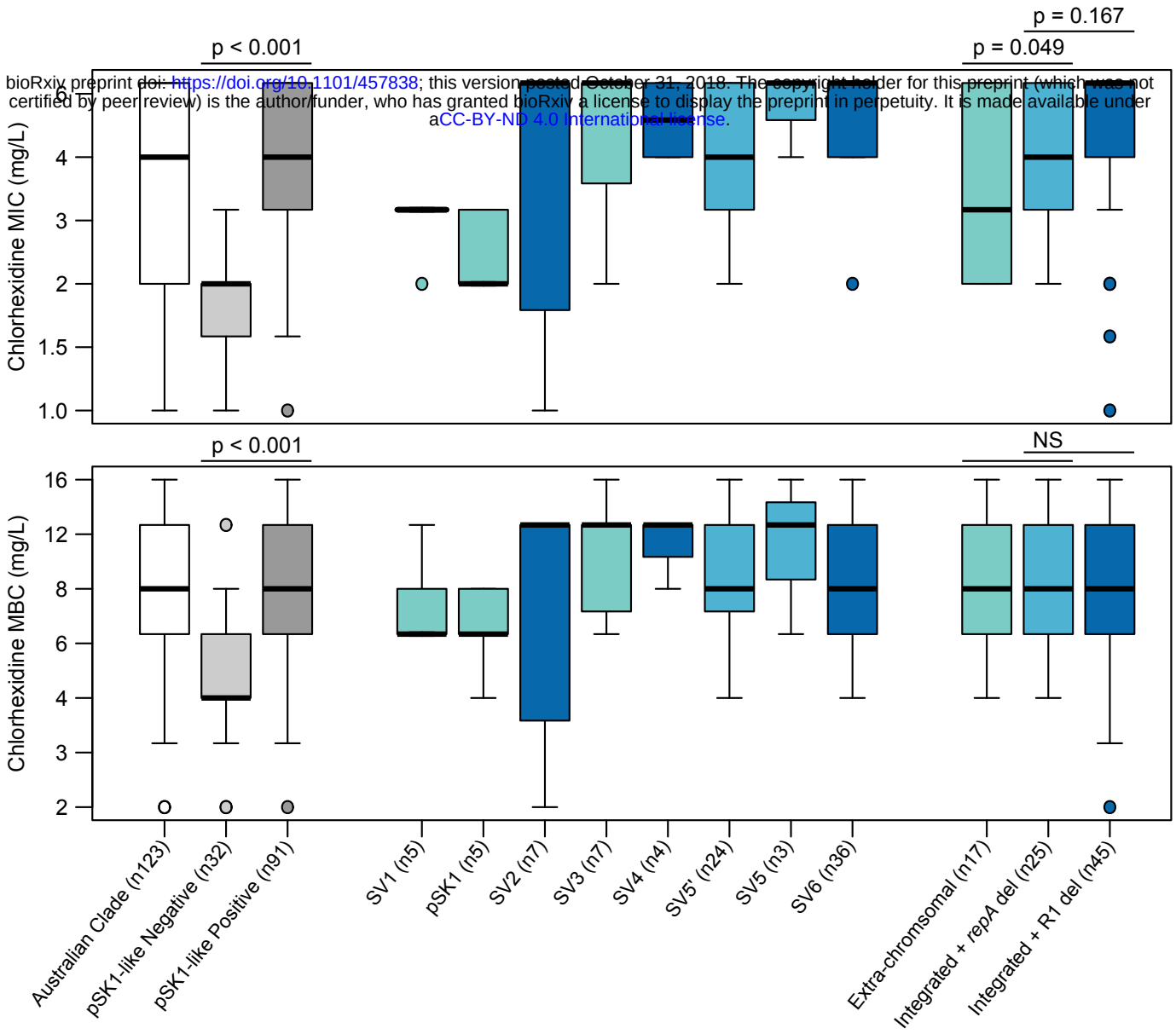

Figure 4. Phenotypic variation in chlorhexidine tolerance. Graphs illustrate the distribution of chlorhexidine MIC (top panel) and MBC (bottom panel) values in the Australian clade. Boxplot features represent the population median (central black line), upper and lower quartiles (box), and range (bars) excluding outliers (circles). Boxplots representing SVs are coloured to reflect a plasmid structural feature: extra-chromosomal plasmid (teal), and chromosomally integrated with either an internal *repA* deletion (light blue) or a multi-CDS plasmid backbone (R1) deletion (dark blue).

Figure 5

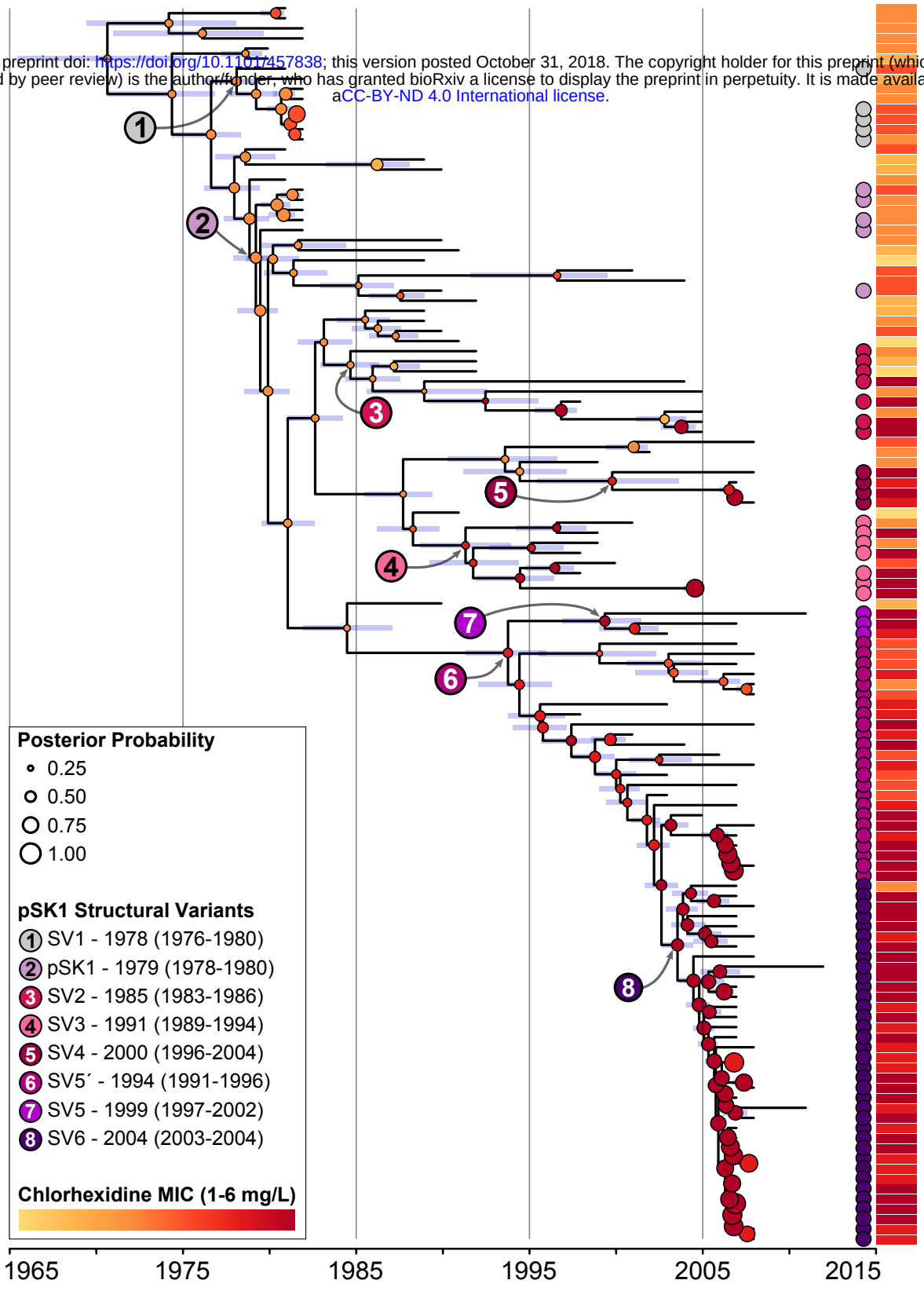

Figure 5. Bayesian phylogenetic model associating chlorhexidine tolerance with pSK1-like plasmid evolution. Illustrated is a maximum clade credibility tree inferred from the whole genome alignment of the Australian clade (n=124). Isolates identified as harbouring a pSK1-like plasmid are indicated by a circle located adjacent to the tree and coloured based on the SV identified. The ancestral nodes in which each SV is estimated to have emerged are indicated by a number (refer to key), those coloured black represent an extra-chromosomal plasmid and those coloured white represent a genomic island. The estimated CHX MIC for all ancestral nodes is indicated by a circle, coloured based on the MIC value and sized according to the posterior probability for the estimate (refer to key). Blue bars represent the 95% highest posterior density interval for the node heights. The aligned heatmap illustrates the phenotypic MIC values attained for each isolate.