Methicillin-resistant *Staphylococcus aureus* genotyping using a small set of polymorphisms

Alex J. Stephens,¹ Flavia Huygens,¹ John Inman-Bamber,¹ Erin P. Price,¹ Graeme R. Nimmo,² Jacqueline Schooneveldt,² Wendy Munckhof³ and Philip M. Giffard¹

Correspondence

Philip M. Giffard

p.giffard@qut.edu.au

¹Cooperative Research Centre for Diagnostics, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland 4001, Australia

²Queensland Health Pathology Services, Princess Alexandra Hospital, Brisbane, Australia

³Infection Management Services, Princess Alexandra Hospital and District Health Service, Brisbane, Australia

The aim of this study was to identify a set of genetic polymorphisms that efficiently divides methicillin-resistant *Staphylococcus aureus* (MRSA) strains into groups consistent with the population structure. The rationale was that such polymorphisms could underpin rapid real-time PCR or low-density array-based methods for monitoring MRSA dissemination in a cost-effective manner. Previously, the authors devised a computerized method for identifying sets of single nucleotide polymorphisms (SNPs) with high resolving power that are defined by multilocus sequence typing (MLST) databases, and also developed a real-time PCR method for interrogating a seven-member SNP set for genotyping *S. aureus*. Here, it is shown that these seven SNPs efficiently resolve the major MRSA lineages and define 27 genotypes. The SNP-based genotypes are consistent with the MRSA population structure as defined by eBURST analysis. The capacity of binary markers to improve resolution was tested using 107 diverse MRSA isolates of Australian origin that encompass nine SNP-based genotypes. The addition of the virulence-associated genes *cna*, *pvl* and *bbp/sdrE*, and the integrated plasmids pT181, pI258 and pUB110, resolved the nine SNP-based genotypes into 21 combinatorial genotypes. Subtyping of the SCCmec locus revealed new SCCmec types and increased the number of combinatorial genotypes to 24. It was concluded that these polymorphisms provide a facile means of assigning MRSA isolates into well-recognized lineages.

Received 17 May 2005

Accepted 21 September 2005

INTRODUCTION

Methicillin-resistant *Staphylococcus aureus* (MRSA) continues to be a significant human pathogen. For many years it has been a common cause of nosocomial infections, and variants capable of causing infections in the community [community-acquired MRSA (CA-MRSA)] are an emerging and serious public-health issue (Chambers, 2001; Diekema *et al.*, 2001). The CA-MRSA phenomenon is remarkable in that these organisms appear to have arisen independently in disparate lineages within the *S. aureus* species (Okuma *et al.*, 2002).

Recent years have seen a rapid increase in the understanding of the comparative genomics and population biology of *S. aureus*, which is leading to an equally rapid increase in the understanding of the relationship between genotype and phenotype (Baba *et al.*, 2002; Robinson & Enright, 2004a). This in turn is facilitating the development of genotyping procedures based upon the interrogation of known genetic polymorphisms (Francois *et al.*, 2004; Robertson *et al.*, 2004; van Leeuwen *et al.*, 2003).

It is now known that the *S. aureus* genome consists of a 'backbone' that is composed of genes that are essentially common to all variants within the species, and which evolves primarily by mutation, punctuated by rare recombination events of unknown mechanism (Baba *et al.*, 2002; Feil *et al.*, 2003; Holden *et al.*, 2004; Kuroda *et al.*, 2001; Robinson & Enright, 2004b). Superimposed upon this is a complement of mobile genes that give rise to binary gene variation between and within lineages (Baba *et al.*, 2002). The staphylococcal cassette chromosome *mec* (SCCmec) is an example of an accessory element whose diversity is able to demonstrate

Abbreviations: CA-MRSA, community-acquired MRSA; *D*, Simpson's index of diversity; MLST, multilocus sequence typing; MRSA, methicillin-resistant *Staphylococcus aureus*; SLV, single locus variant; SNP, single nucleotide polymorphism; ST, sequence type.

A complete description of the MRSA isolates used in this study is available in Supplementary Table S1 with the online version of this paper.

the multiple insertion events that have occurred in the evolution of MRSA (Ito *et al.*, 2001; Katayama *et al.*, 2000).

Our overarching aim is to develop microbial genotyping methods based upon small sets of polymorphisms selected from known genomic diversity on the basis of their optimal combinatorial informative power. Such methods would be suitable for low-density array, or real-time PCR technology platforms, and would facilitate routine surveillance, infection control, and rapid and cost-effective identification of clones of interest. We have previously reported a computerized method for identifying sets of single nucleotide polymorphisms (SNPs) in multilocus sequence typing (MLST) databases that provide high Simpson's index of diversity (*D*) values with respect to the databases, and the application of this to *S. aureus* (Robertson *et al.*, 2004). A set of seven SNPs was identified that provides a *D* of 0.95 with respect to the *S. aureus* MLST database, and a real-time PCR method for interrogating these SNPs was reduced to practice. Robertson and co-workers did not determine the concordance between SNP profiles and the *S. aureus* population structure, i.e. whether the SNP-based genotyping clusters closely related isolates and discriminates more distantly related ones (Robertson *et al.*, 2004). Accordingly, one of the principal aims of this study was to address this question as it relates to MRSA, and also to apply the SNP-based genotyping method to a larger collection of MRSA isolates.

An ideal molecular typing method would indicate the position of the isolate within the species population structure, and also economically and conveniently serve as a high-resolution comparative method. The 'progressive hierarchical resolving assays using nucleic acids' (PHRANA) concept articulated by Keim *et al.* (2004) is very helpful in devising such typing methods. The underlying principle of PHRANA is that slowly evolving polymorphic sites do not provide high resolution, while rapidly evolving markers are unreliable at indicating the position of the isolate within the population structure and are subject to homoplasy if used by themselves. However, rapidly and slowly evolving markers in combination provide a usefully complete picture. Therefore, the other principal aim of this study was to identify a small set of binary markers that adds resolving power to the SNP set. The markers identified are four virulence-associated genes and three SCCmec-associated integrated plasmids. Polymorphisms in SCCmec and the *agr* locus have previously been used in combination with MLST to define MRSA clones (Oliveira *et al.*, 2001), so we have also assessed the informative powers of SCCmec and *agr* typing in this context. The genotypes defined by the combination of SNPs and binary markers are very easy to obtain and digitize, consistent with previous studies, and provide easily understood genetic fingerprints for the well-known MRSA clones.

METHODS

Bacterial isolates. One hundred and seven MRSA isolates from four different collections were included in this study. Forty-seven isolates were from the Australian Group on Antimicrobial Resistance

(AGAR) 2000 community-onset collection (Coombs *et al.*, 2004), 15 isolates were from a cluster of community-onset disease in South-East Queensland, Australia, nine were from another similar cluster of community-onset disease, and 36 isolates have been previously described (Huygens *et al.*, 2004; Nimmo *et al.*, 2000; Peleg *et al.*, 2005). A complete description of the isolates is available in Supplementary Table S1.

Cultivation and DNA template extraction. Each isolate was grown on brain heart infusion (BHI) agar plates overnight at 37 °C. A single colony was selected and grown in 5 ml BHI broth overnight in an incubating shaker at 37 °C. The DNA template was extracted using the QIAGEN DNA mini kit, as recommended by the manufacturer, with 200 µg lysostaphin ml⁻¹ (Sigma). The DNA template was quantified using UV spectrophotometry before use in PCR.

MLST clonal group determination using SNP-based genotyping. Allele-specific real-time PCR and the clustering algorithm eBURST were used to group uncharacterized isolates into MLST clonal complexes. The SNP-based genotyping methodology has been previously described (Robertson *et al.*, 2004); however, the total reaction volume was reduced from 25 to 20 µl.

Binary gene selection, primer design and PCR amplification. Virulence genes that were candidates for intra-MLST clonal complex variation were selected based on previous studies (Moore & Lindsay, 2001; Peacock *et al.*, 2002). Real-time PCR primers for the genes *fnbA*, *cna*, *icaA*, *sdrE*, *bbp*, *hlg*, *pvl* and *tst* (Table 1) were designed using the Primer Express software package (Applied Biosystems). Primer sequences and PCR conditions for amplifying the plasmids pT181, pUB110 and pI258 have been published previously (Huygens *et al.*, 2002). Virulence gene amplification was conducted on an ABI7000 real-time PCR device with the following PCR cycling conditions: stage 1, 50 °C for 2:00 (MM:SS); stage 2, 95 °C for 10:00; stage 3 step 1, 95 °C for 00:15; stage 3 step 2, 60 °C for 01:00; repeat stage 3 30 times. Each 20 µl reaction was conducted in duplicate and contained 10 µl 2× SYBR green PCR master mix (Applied Biosystems), 1 µl of a 2.5 µM solution of each forward and reverse primer, and 8 µl DNA template containing approximately 10 ng DNA.

MLST. MLST was conducted as described elsewhere (Enright *et al.*, 2000). Twenty-four isolates from this study have previously been genotyped using MLST (Coombs *et al.*, 2004).

SCCmec typing. SCCmec element type was determined by PCR identification of the *ccr* (cassette chromosome recombinase) allele and the *mec* complex class, as described elsewhere (Okuma *et al.*, 2002). The recently characterized *ccrC* from SCCmec V was identified using primers described by Ito *et al.* (2004). Isolates previously characterized for SCCmec are described by Coombs *et al.* (2004).

agr genotyping. The *agr* locus alleles were amplified using the multiplex PCR procedure described by Lina *et al.* (2003).

RESULTS AND DISCUSSION

Informative power of a seven-member SNP set for genotyping MRSA

We have previously reported the derivation from the *S. aureus* MLST database of a highly discriminatory set of seven SNPs. The identification of the SNPs was aided by the computer program Minimum SNPs, which can take as input a complete MLST database and provide as output sets of SNPs that, if interrogated, provide a high Simpson's index of

Table 1. Primer sequences

Virulence gene	Gene function	GenBank accession no.	Primer	Sequence (5'–3')
<i>fnbA</i>	Fibronectin binding protein	J04151	<i>fnbA</i> _F1	ACA GTA ACA GAA CAA CCG TCA AAC G
			<i>fnbA</i> _R1	TTG CTG GTT GTG CAG TTT GTG
<i>cna</i>	Collagen adhesin protein	M81736	<i>cna</i> _F1	TCA AGC AGT TAT TAC ACC AGA CGG
			<i>cna</i> _R1	GCA AAT CCC GAA ACA TCA CTT AA
<i>sdrE</i>	Putative adhesin	AJ005647	<i>sdrE</i> _F2	CAT CAG ATG GCG AAC TAG ATA TTG C
			<i>sdrE</i> _R2	CCA GTG TCA TTA GAA GTT ACG ATG AAG
<i>bbp</i>	Bone binding protein	Y18653	<i>bbp</i> _F1	AAT GAC AGT AAA GTA TGA CAA GGA GAC G
			<i>bbp</i> _R1	TTC CCT TCG CTA TTA CTT CAC CTG
<i>icaA</i>	Intracellular adhesin	AF086783	<i>icaA</i> _F1	TCG ACG TTG GCT ACT GGG AT
			<i>icaA</i> _R1	GGC AAG CGG TTC ATA CTT AAT ACG
<i>pvl</i>	Leukocidin	X72700	<i>pvl</i> _F1	TCA CAA AAT GCC AGT GTT ATC CAG
			<i>pvl</i> _R1	TTT TGC AGC GTT TTG TTT TCG
<i>hlg</i>	Haemolysin	L01055	<i>hlg</i> _F1	ATG AAA TCA AGG TGA AAG GAC AGA A
			<i>hlg</i> _R1	GAC TTT TTT TAC GCT GAC TGG TGT
<i>tst</i>	Toxic shock syndrome toxin	AY074881	<i>tst</i> _F1	CCC TTT GTT GCT TGC GAC A
			<i>tst</i> _R1	CCC ACT ACT ATA CCA GTC TAG CAA ATC C

diversity, D , with respect to the database. We term this a 'high D ' SNP set.

It is desirable that a genotyping method provides outputs that are consistent with the population structure of the species in question (i.e. similarity of genotypes is correlated with relationship of isolates), as this facilitates the inference of dissemination and evolutionary events. Accordingly, the consistency between the MRSA population structure and the SNP profiles was determined (Fig. 1). The population structure has been depicted using an eBURST analysis of MRSA sequence types present in the *S. aureus* MLST database as at January 25 2005. eBURST was used because conventional phylogenetic analyses of MLST databases may be confounded by recombination events. The eBURST algorithm identifies clonal complexes using MLST allele data and attempts to identify clonal complex progenitors (Feil *et al.*, 2004). This analysis showed that seven SNPs resolve known human-derived MRSA sequence types (STs) into 29 SNP profiles. The major hospital-acquired and community-acquired clones are discriminated, with the exception of the unrelated STs 59 and 93, which represent the North American SF25 : P (Diep *et al.*, 2004) and Queensland clones (Munckhof *et al.*, 2003) of CA-MRSA, respectively. Even though the complete set of STs includes many that were not known at the time that the SNP set was defined, the SNP genotypes are concordant with the population structure.

It is obvious that a small SNP set cannot have the discriminatory power of full ST determination. In Fig. 1, it can be seen that there are two related but different phenomena. Firstly, there are instances in which the SNPs fail to discriminate unrelated singletons or clonal complexes. We have found that increasing the size of the SNP set from seven to 14 confers the ability to resolve the great majority of unrelated clonal complexes and singletons (data not shown).

It is reasonable to state that the optimum size of an SNP set is a function of the need to discriminate singletons and small clonal complexes. We have chosen to use seven SNPs, as this discriminates the major clones and also indicates whether or not an isolate belongs to a more unusual lineage. The addition of another seven SNPs would reveal which minor clonal complex or singleton an unusual isolate belonged to.

Secondly, the SNP set provides very little power to discriminate between members of clonal complexes. This is expected, as *S. aureus* is a clonal organism with low diversity, and the great majority of single locus variants (SLVs) of clonal complex progenitors have arisen by mutation (Feil *et al.*, 2003). For a given SLV, the SNP that discriminates it from the clonal complex progenitor is not polymorphic anywhere else in the species, and so will not be readily identified in a computerized search for high D SNP sets. An SNP-based method for resolving within *S. aureus* clonal complexes as defined by MLST would need to individually resolve each clonal complex member from the clonal complex progenitor, and so would be most inefficient. Consistent with this, in the majority of the small number of instances in which the SNP profile of an SLV differs from the SNP profile of the clonal complex progenitor, the SLV has arisen by recombination rather than mutation. For example, the alleles that discriminate STs 221, 194 and 148 from ST5 are found elsewhere in the MLST database, while the corresponding alleles for the other SLVs of ST5 are unique. ST221 is particularly interesting in that it has acquired an allele that is characteristic of ST1, and ST221 shares four of the seven alleles with ST1, as well as an SNP profile. In this instance, the analysis of the SNP profiles has revealed a recombination event.

It is unclear how informative intra-clonal complex population structures as revealed by eBURST analysis are,

Fig. 1. Output of an eBURST analysis of *S. aureus* ST in the MLST database listed as being MRSA of human origin. The group definition was set to 'zero alleles in common' in order to allow visualization of all STs on a single diagram. Solid lines represent single locus variation, and double locus variation is not indicated. Each circle represents a particular SNP profile. The SNP profiles are indicated in the order *arcC210*, *tpi243*, *arcC162*, *tpi241*, *yqiL333*, *aroE132*, *gmk129*. The diameters of the solid spots are proportional to the occurrences of the STs in the MLST database, i.e. a large spot means that a large number of isolates with that ST have been entered into the database. SNP profiles found in this study are boxed, together with the number of isolates.

SNP-based genotyping of Australian MRSA

2004), a selection of MRSA isolates of Australian origin were subject to SNP typing. Firstly, the SNP interrogation method was validated using 32 MRSA isolates of known ST. The kinetic PCR was completely successful at calling the bases at the SNPs (see Supplementary Table S1). For every polymorph (allelic state of an SNP), the reaction with the perfectly matched primer reached the threshold in fewer cycles than the reaction with the mismatched primer, i.e. the ΔC_T was always in the expected orientation. It was concluded that the kinetic PCR approach to interrogating these SNPs is robust. The 11 STs (1, 8, 22, 30, 45, 73, 78, 88, 93, 128 and 239) were resolved into nine SNP profiles. The unresolved STs (ST239–ST128 and ST78–ST88) are very closely related to each other (Supplementary Table S1, Fig. 1).

Table 2. SNP profiles for isolates of unknown ST

The figures in parentheses in the 'ST' column are the STs found in the isolates of known ST (see Supplementary Table S1) that are consistent with the SNP profile.

Isolate	ST	PFGE type	<i>arcC210C/T/A</i>	<i>tpi243A/G</i>	<i>arcC162T/A</i>	<i>tpi241G/A</i>	<i>yqiL333C/T</i>	<i>aroE132A/G</i>	<i>gmkl29C/T</i>
D828570	(30)	A0	T	G	A	G	T	G	T
E822547	(30)	A0	T	G	A	G	T	G	T
F810539	(30)	A0	T	G	A	G	T	G	T
MB826559	(30)	A0	T	G	A	G	T	G	T
E803534	(30)	A0	T	G	A	G	T	G	T
MD817541	(30)	A0	T	G	A	G	T	G	T
MA830538	(30)	A0	T	G	A	G	T	G	T
I825560	(30)	A0	T	G	A	G	T	G	T
MA823547	(30)	A1	T	G	A	G	T	G	T
MC810534	(30)	A1	T	G	A	G	T	G	T
MD821552	(30)	A2	T	G	A	G	T	G	T
E802537	(30)	A3	T	G	A	G	T	G	T
68284/98	(30)	A5	T	G	A	G	T	G	T
E822485	(30)	B	T	G	A	G	T	G	T
MJ710566	(45)	C	C	G	A	G	T	G	T
B8-31	(45)	K	C	G	A	G	T	G	T
MD828534	(1)	E	C	A	T	G	C	A	C
MB827549	(1)	E	C	A	T	G	C	A	C
MD808118	(1)	L	C	A	T	G	C	A	C
K704540	(8)	F	T	A	A	G	T	A	C
K705613	(8)	F2	T	A	A	G	T	A	C
K711532	(8)	F3	T	A	A	G	T	A	C
K714372	(8)	F4	T	A	A	G	T	A	C
MIC703484	(8)	G1	T	A	A	G	T	A	C
I823541	(8)	G2	T	A	A	G	T	A	C
E812560	(8)	J	T	A	A	G	T	A	C
IPOOM14235	(8)	O	T	A	A	G	T	A	C
E804531	(5)	I	C	A	T	G	C	A	T
IP01M1081	(93)	R	T	G	T	G	T	A	T
IP01M2046	(78, 88)	P1	T	A	T	G	C	A	T

ST93 and ST239 (ST73 was found in the fully sequence-typed isolates, and is an SLV of ST5 which is a very well known, and common, clonal complex progenitor). The SNP profiles showed consistency with the PFGE profiles and origins of the isolates. The ST30/pulsotype A isolates are members of the Western Samoan Phage Pattern (WSPP)/South West Pacific (SWP) clone that is strongly associated with community-acquired infections in Oceania (Nimmo *et al.*, 2000). As expected, PFGE provided higher resolution than the SNPs, and on no occasion were different SNP profiles associated with the same pulsotype. The STs included in this study encompass much of the diversity of Australian MRSA (Coombs *et al.*, 2004). It was concluded that the high *D* SNPs are useful for identifying the major Australian MRSA clones.

Identification of a set of genes that exhibits useful binary variability

The high *D* SNP set is effective at deciphering the stable genome backbone and characterizing isolates into the MLST

clonal complexes. In order to efficiently add resolution, we set out to identify a small set of genes that exhibits complementary binary variability, i.e. their presence or absence is not predictable either from the SNP profile or from the presence or absence of other binary markers. Ideally, the markers would also have the potential to directly provide clinically significant information. Peacock *et al.* (2002) identified a set of virulence genes whose distribution between and within specific *S. aureus* lineages did not correlate with ST, and therefore were seen as ideal candidates. These genes were *pvl*, *cna*, *sdrE/bbp* alleles, *tst*, *icaA*, *fnbA* and *hlg*.

In order to identify the gene variation within our collection, real-time PCR assays were designed and optimized. The PCR amplicon length for each reaction was verified using agarose gel electrophoresis. The results from the PCR assays indicated that the *pvl*, *cna*, *sdrE/bbp* and *tst* genes displayed binary variability within the collection, and that the *icaA*, *fnbA* and *hlg* genes were present in all isolates (Table 3). The distributions of the variable virulence genes are consistent with previous reports (Moore & Lindsay, 2001; Peacock

Table 3. Gene prevalence

Gene	Isolate acquisition		Total prevalence (%)
	Community (<i>n</i> =66) (%)	Nosocomial (<i>n</i> =41) (%)	
<i>fnbA</i>	66 (100)	41 (100)	107 (100)
<i>cna</i>	54 (81.8)	35 (85.4)	89 (83.2)
<i>sdrE</i>	35 (53)	39 (95.1)	74 (69.2)
<i>bbp</i>	29 (43.9)	1 (2.4)	30 (28)
<i>icaA</i>	66 (100)	41 (100)	107 (100)
<i>pvl</i>	48 (72.7)	0 (0)	48 (44.9)
<i>hlg</i>	66 (100)	41 (100)	107 (100)
<i>tst</i>	0 (0)	4 (9.8)	4 (3.7)

et al., 2002). The *pvl* and *tst* genes are known to be mobile, the *sdrE* locus is polymorphic, and *cna* has been reported not to be consistently carried throughout the species, although the basis for its mobility is not understood. The three genes common to all isolates are chromosomally located and do not have any obvious means of mobility, although variation of these genes has been demonstrated elsewhere (Peacock *et al.*, 2002). The *tst* gene was found in only a very small number of isolates, so in the context of this study was shown to have limited usefulness. Therefore, the informative power of *pvl*, *cna* and *sdrE/bbp* in combination with the high *D* SNPs was assessed.

In the case of the CA-MRSA isolates, the ST1, ST30 and ST93 SNP profiles were divided into a total of three, two and two subtypes, respectively. Interestingly, each CA-MRSA-associated SNP profile was found to encompass both *pvl*-positive and -negative isolates. ST36 is an SLV of ST30, and is not associated with community acquisition. The single ST36 isolate in our collection was differentiated from the CA-MRSA ST30 strain by being *pvl* negative and *tst* positive. The ST93 SNP profile encompassed two PFGE pulsotypes, Q and R. The ST93, R pulsotype is known as the CA-MRSA Queensland clone, and is characteristically *pvl* positive. Just one isolate was pulsotype Q and was *pvl* negative; full MLST determination revealed the isolate as ST59. Analysis using eBURST indicates that ST59 and ST93 are not closely related, and coincidentally share the same SNP profile. This is not fortuitous, as the emerging CA-MRSA SF25:P clone in the United States is ST59 (Diep *et al.*, 2004). Although the ST59 isolate in this study is *pvl* negative, and so could be discriminated from the ST93 isolates on that basis, ST59 can be *pvl* positive (Diep *et al.*, 2004) and therefore difficult to discriminate from the Queensland clone. However, because ST93 is a singleton, we were able to use minimum SNPs to determine that ST93 can be discriminated from all other STs in the database by the presence of e.g. a G at *aroE*252. This SNP could easily be added to a standardized procedure based on the high *D* SNP set, or alternatively could underpin a separate single-SNP assay that is only carried out when the high *D* SNP profile TGTGTAT is obtained.

With respect to the nosocomial isolates, the ST239 group consisted of 24 isolates, but was only divided into two subtypes based on virulence gene variation; a single isolate tested negative to both *sdrE* and *bbp*. The two isolates of the ST5 SNP profile were found to differ at the *cna* locus, the ST5 isolate was negative and the ST73 isolate was positive. The three ST8 and three ST78 SNP group isolates were also divided through the lack of the *cna* gene in two isolates from each group. Variation of the virulence gene targets within the SNP groups of ST22 and 45 was not identified, most likely due to the small sizes of these groups.

SCCmec-associated plasmids add resolving power

It is well known that one mode of variation in the SCCmec element is the plasmids integrated at the 3' end (Oliveira *et al.*, 2000). Our group has previously shown that these plasmids can be used for binary typing (Huygens *et al.*, 2002). Here, we have determined the resolving power of these integrated plasmids in combination with the high *D* SNPs and the variable virulence genes.

The entire collection was tested for these plasmids. Variation was identified in the ST1, ST30 and ST239 SNP profiles. An additional subtype was identified from the CA-MRSA-associated ST1 SNP profile, with five *pvl*-negative isolates testing positive for pT181. The ST36 isolate sharing the ST30 SNP profile was further differentiated from the ST30 strain by testing positive for pUB110. Most variation was in the ST239 SNP profile isolates, among which four plasmid profiles were found. The majority carried pT181 and pI258, with the remaining isolates carrying either pI258, pUB110 or no plasmids. It was concluded that these plasmids efficiently increased genotyping resolution, particularly with nosocomial isolates, which characteristically harbour the large forms of the SCCmec element.

Comparison of combinatorial genotyping with SCCmec and *agr* typing

One combinatorial genotyping method that has become well accepted is SCCmec typing in combination with ST determination. *mec* typing is a binary system based upon variation in the *ccr* gene complex and the upstream *mecA* regulatory region (Okuma *et al.*, 2002). It was therefore of interest to compare the informative power of *mec* typing with that of the other binary markers.

SCCmec characterization was undertaken on 83 isolates. Each of the five major structural types was successfully identified (SCCmec I–V), along with at least four novel variants. The five ST1 SNP profile isolates that tested negative for *pvl* and positive for pT181 were found to carry a novel SCCmec element composed of *ccrA*1 with *mec* class C. Other novel structural variants were identified in two unrelated isolates, which tested positive for *mec* class B and negative for each *ccr* allele. Surprisingly, 10 of the 29 ST30 SCCmec type IV isolates tested negative for subtype a, as did

Table 4. SNP and binary gene combinatorial genotypes obtained in this study

The SNP profiles are shown in column 1, in the order *arcC210*, *tpi243*, *arcC162*, *tpi241*, *yqiL333*, *aroE132*, *gmk129*. The numbers represent the clonal complex. The 'Clone' column refers to the common names for these lineages.

SNP genotype	No. of isolates	Profile								Clone		Reference	
		<i>pvl</i>	<i>cna</i>	<i>sdrE/bbp</i>	<i>tst</i>	pT181	pI258	pUB110	SCC <i>mec</i>	<i>agr</i>			
CATGCAC													
1a	3	+	+	<i>sdrE</i>	—	—	—	—	IV	3	WAMRSA-1	Coombs <i>et al.</i> (2004)	
1b	11	—	+	<i>sdrE</i>	—	—	—	—	IV, I	3	WAMRSA-1	Coombs <i>et al.</i> (2004)	
1c	5	—	+	<i>sdrE</i>	—	+	—	—	New	3			
1d	2	—	+	Neither	—	—	—	—	IV	3			
CATGCAT													
5a	1	—	—	<i>sdrE</i>	—	—	—	—	IV	2			
5b	1	—	+	<i>sdrE</i>	—	—	—	—	IV	2			
TATGTAC													
8a	2	—	—	<i>sdrE</i>	—	—	—	—	IV, V	1	nmMRSAD	Coombs <i>et al.</i> (2004)	
8b	1	—	+	<i>sdrE</i>	—	—	—	—	New	1	Irish-2	Coombs <i>et al.</i> (2004)	
CGTGTGC													
22	4	—	+	<i>sdrE</i>	—	—	—	—	IV	1	EMRSA-15	Coombs <i>et al.</i> (2004)	
TGAGTGT													
30a	29	+	+	<i>bbp</i>	—	—	—	—	IV	3	SWP/WSPP	Coombs <i>et al.</i> (2004)	
30b	1	—	+	<i>bbp</i>	+	—	—	+	II	3	EMRSA-16	Nimmo <i>et al.</i> (2000)	
CGAGTGT													
45	3	—	+	<i>sdrE</i>	+	—	—	—	V	1	nmMRSAE	Coombs <i>et al.</i> (2004)	
TATGCAT													
78a	2	—	—	<i>sdrE</i>	—	—	—	—	IV, New	3			
78b	1	—	+	<i>sdrE</i>	—	—	—	—	IV	3			
TGTGTAT													
93a	16	+	—	<i>sdrE</i>	—	—	—	—	IV	3	QLD	Coombs <i>et al.</i> (2004)	
93b	1	—	—	<i>sdrE</i>	—	—	—	—	IV	1			
TAAGTAC													
239a	10	—	+	<i>sdrE</i>	—	+	+	—	III	1	AUS-3	Coombs <i>et al.</i> (2004)	
239b	8	—	+	<i>sdrE</i>	—	—	—	+	III	1	AUS-2	Coombs <i>et al.</i> (2004)	
239c	4	—	+	<i>sdrE</i>	—	—	+	—	III	1			
239d	1	—	+	<i>sdrE</i>	—	—	—	—	III	1			
239e	1	—	+	Neither	—	+	+	—	III	1	AUS-3	Coombs <i>et al.</i> (2004)	

the ST22 SNP group. These isolates were further analysed and proved negative for subtypes b and c. The ST239 SNP profile isolates tested in this study were each found to carry SCC*mec* type III, but as stated in the previous section, these elements vary with respect to integrated plasmids that define the subgroups. There was also a single ST1 SNP isolate (FH53) that carried SCC*mec* type I. In conclusion, SNP genotyping used in combination with SCC*mec* types, including the novel variants, defined 17 genotypes. This is fewer than the 21 genotypes defined by the SNPs+binary virulence genes+integrated plasmids, although 24 genotypes were defined when all loci were interrogated. SCC*mec* subtyping could define more genotypes, but the subtypes have not been fully defined, with some isolates remaining un-subtypable. Also, the SCC*mec* elements are very diverse and appear to evolve very rapidly, thus making it difficult to design a generally applicable and easily interpretable typing method. It may be that the SCC*mec* element will prove most valuable for providing higher genotyping resolution on occasions when this is required.

The *agr* locus is a polymorphic region of the genome that modulates expression of a set of virulence-related genes (Ji *et al.*, 1995). To determine if variation at this locus is useful for adding resolution, the collection was assayed for each of the four *agr* types. A single instance of intra-SNP profile variation at the *agr* locus was identified. The *pvl*-negative ST59 isolate that shares an SNP profile with the ST93 SNP genotype was found to have *agr* type I, compared to the ST93 isolates that carry *agr* type III. These results indicate that with this collection of isolates, binary variation is concordant with the genotypes defined by the SNPs and the other binary markers, so interrogation of the *agr* locus adds little informative power in this context.

The genotypes obtained in this study are collated in Table 4.

Conclusion

A combination of seven SNPs and seven binary targets resolves a diverse collection of Australian MRSA isolates

into 21 genotypes that are unambiguous, easily obtainable and fully consistent with previous studies. Interrogation of polymorphisms in the SCC mec and *agr* loci has the potential to increase resolution in certain circumstances, as does the addition of more SNPs.

ACKNOWLEDGEMENTS

This work was funded by the Australian Federal Government Cooperative Research Centres Program. The authors thank the Australian Group for Antimicrobial Resistance for the provision of isolates.

REFERENCES

- Baba, T., Takeuchi, F., Kuroda, M. & 11 other authors (2002). Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359**, 1819–1827.
- Chambers, H. F. (2001). The changing epidemiology of *Staphylococcus aureus*? *Emerg Infect Dis* **7**, 178–182.
- Coombs, G. W., Nimmo, G. R., Bell, J. M., Huygens, F., O'Brien, F. G., Malkowski, M. J., Pearson, J. C., Stephens, A. J. & Giffard, P. M. (2004). Genetic diversity among community methicillin-resistant *Staphylococcus aureus* strains causing outpatient infections in Australia. *J Clin Microbiol* **42**, 4735–4743.
- Diekema, D. J., Pfaller, M. A., Schmitz, F. J., Smayevsky, J., Bell, J., Jones, R. N. & Beach, M. (2001). Survey of infections due to *Staphylococcus species*: frequency of occurrence and antimicrobial susceptibility of isolates collected in the United States, Canada, Latin America, Europe, and the Western Pacific region for the SENTRY Antimicrobial Surveillance Program, 1997–1999. *Clin Infect Dis* **32**, S114–S132.
- Diep, B. A., Sensabaugh, G. F., Somboona, N. S., Carleton, H. A. & Perdreau-Remington, F. (2004). Widespread skin and soft-tissue infections due to two methicillin-resistant *Staphylococcus aureus* strains harboring the genes for Pantone–Valentine leucocidin. *J Clin Microbiol* **42**, 2080–2084.
- Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. (2000). Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol* **38**, 1008–1015.
- Feil, E. J., Cooper, J. E., Grundmann, H. & 9 other authors (2003). How clonal is *Staphylococcus aureus*? *J Bacteriol* **185**, 3307–3316.
- Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P. & Spratt, B. G. (2004). eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* **186**, 1518–1530.
- Francois, P., Renzi, G., Pittet, D., Bento, M., Lew, D., Harbarth, S., Vaudaux, P. & Schrenzel, J. (2004). A novel multiplex real-time PCR assay for rapid typing of major staphylococcal cassette chromosome *mec* elements. *J Clin Microbiol* **42**, 3309–3312.
- Holden, M. T., Feil, E. J., Lindsay, J. A. & 42 other authors (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.
- Huygens, F., Nimmo, G. R., Schooneveldt, J., Munckhof, W. J. & Giffard, P. M. (2002). Genotyping of methicillin-resistant *Staphylococcus aureus* by assaying for the presence of variable elements associated with *mecA*. *J Clin Microbiol* **40**, 3093–3097.
- Huygens, F., Stephens, A. J., Nimmo, G. R. & Giffard, P. M. (2004). *mecA* locus diversity in methicillin-resistant *Staphylococcus aureus* isolates in Brisbane, Australia, and the development of a novel diagnostic procedure for the Western Samoan phage pattern clone. *J Clin Microbiol* **42**, 1947–1955.
- 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.
- Ito, T., Ma, X. X., 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**, 2637–2651.
- Ji, G., Beavis, R. C. & Novick, R. P. (1995). Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A* **92**, 12055–12059.
- Katayama, Y., Ito, T. & Hiramatsu, K. (2000). A new class of genetic element, staphylococcal cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **44**, 1549–1555.
- Keim, P., Van Ert, E. M., Pearson, T., Vogler, A. J., Huynh, L. Y. & Wagner, D. M. (2004). Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infection, Genetics and Evolution* **4**, 205–213.
- Kuroda, M., Ohta, T., Uchiyama, I. & 34 other authors (2001). Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**, 1225–1240.
- Lina, G., Boutite, F., Tristan, A., Bes, M., Etienne, J. & Vandenesch, F. (2003). Bacterial competition for human nasal cavity colonization: role of *Staphylococcal agr* alleles. *Appl Environ Microbiol* **69**, 18–23.
- Moore, P. C. & Lindsay, J. A. (2001). Genetic variation among hospital isolates of methicillin-sensitive *Staphylococcus aureus*: evidence for horizontal transfer of virulence genes. *J Clin Microbiol* **39**, 2760–2767.
- Munckhof, W. J., Schooneveldt, J., Coombs, G. W., Hoare, J. & Nimmo, G. R. (2003). Emergence of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) infection in Queensland, Australia. *Int J Infect Dis* **7**, 259–264.
- Nimmo, G. R., Schooneveldt, J., O'Kane, G., McCall, B. & Vickery, A. (2000). Community acquisition of gentamicin-sensitive methicillin-resistant *Staphylococcus aureus* in southeast Queensland, Australia. *J Clin Microbiol* **38**, 3926–3931.
- Okuma, K., Iwakawa, K., Turnidge, J. D. & 10 other authors (2002). Dissemination of new methicillin-resistant *Staphylococcus aureus* clones in the community. *J Clin Microbiol* **40**, 4289–4294.
- Oliveira, D. C., Wu, S. W. & de Lencastre, H. (2000). Genetic organization of the downstream region of the *mecA* element in methicillin-resistant *Staphylococcus aureus* isolates carrying different polymorphisms of this region. *Antimicrob Agents Chemother* **44**, 1906–1910.
- Oliveira, D. C., Tomasz, A. & de Lencastre, H. (2001). The evolution of pandemic clones of methicillin-resistant *Staphylococcus aureus*: identification of two ancestral genetic backgrounds and the associated *mec* elements. *Microb Drug Resist* **7**, 349–361.
- Peacock, S. J., Moore, C. E., Justice, A., Kantzanou, M., Story, L., Mackie, K., O'Neill, G. & Day, N. P. (2002). Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. *Infect Immun* **70**, 4987–4996.
- Peleg, A. Y., Munckhof, W. J., Kleinschmidt, S. L., Stephens, A. J. & Huygens, F. (2005). Life-threatening community-acquired methicillin-resistant *Staphylococcus aureus* infection in Australia. *Eur J Clin Microbiol Infect Dis* **24**, 384–387.
- Robertson, G. A., Thiruvankataswamy, V., Shilling, H., Price, E. P., Huygens, F., Henskens, F. A. & Giffard, P. M. (2004). Identification

and interrogation of highly informative single nucleotide polymorphism sets defined by bacterial multilocus sequence typing databases. *J Med Microbiol* **53**, 35–45.

Robinson, D. A. & Enright, M. C. (2004a). Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. *Clin Microbiol Infect* **10**, 92–97.

Robinson, D. A. & Enright, M. C. (2004b). Evolution of *Staphylococcus aureus* by large chromosomal replacements. *J Bacteriol* **186**, 1060–1064.

Robinson, D. A., Kearns, A. M., Holmes, A. & 8 other authors (2005). Re-emergence of early pandemic *Staphylococcus aureus* as a community-acquired methicillin-resistant clone. *Lancet* **365**, 1256–1258.

van Leeuwen, W. B., Jay, C., Snijders, S., Durin, N., Lacroix, B., Verbrugh, H. A., Enright, M. C., Troesch, A. & van Belkum, A. (2003). Multilocus sequence typing of *Staphylococcus aureus* with DNA array technology. *J Clin Microbiol* **41**, 3323–3326.