

RESEARCH ARTICLE

Open Access

Single-molecule sequencing reveals the molecular basis of multidrug-resistance in ST772 methicillin-resistant *Staphylococcus aureus*

Eike J Steinig¹, Patiyan Andersson¹, Simon R Harris², Derek S Sarovich¹, Anand Manoharan³, Paul Coupland², Matthew TG Holden⁴, Julian Parkhill², Stephen D Bentley², D Ashley Robinson⁵ and Steven YC Tong^{1*}

Abstract

Background: Methicillin-resistant *Staphylococcus aureus* (MRSA) is a major cause of hospital-associated infection, but there is growing awareness of the emergence of multidrug-resistant lineages in community settings around the world. One such lineage is ST772-MRSA-V, which has disseminated globally and is increasingly prevalent in India. Here, we present the complete genome sequence of DAR4145, a strain of the ST772-MRSA-V lineage from India, and investigate its genomic characteristics in regards to antibiotic resistance and virulence factors.

Results: Sequencing using single-molecule real-time technology resulted in the assembly of a single continuous chromosomal sequence, which was error-corrected, annotated and compared to nine draft genome assemblies of ST772-MRSA-V from Australia, Malaysia and India. We discovered numerous and redundant resistance genes associated with mobile genetic elements (MGEs) and known core genome mutations that explain the highly antibiotic resistant phenotype of DAR4145. Staphylococcal toxins and superantigens, including the leukotoxin Panton-Valentini Leukocidin, were predominantly associated with genomic islands and the phage ϕ -IND772PVL. Some of these mobile resistance and virulence factors were variably present in other strains of the ST772-MRSA-V lineage.

Conclusions: The genomic characteristics presented here emphasize the contribution of MGEs to the emergence of multidrug-resistant and highly virulent strains of community-associated MRSA. Antibiotic resistance was further augmented by chromosomal mutations and redundancy of resistance genes. The complete genome of DAR4145 provides a valuable resource for future investigations into the global dissemination and phylogeography of ST772-MRSA-V.

Keywords: *Staphylococcus aureus*, MRSA, ST772, Antibiotic resistance, Mobile genetic elements, Complete genome, DAR4145, India

Background

Staphylococcus aureus is one of the leading causes of hospital-associated infections worldwide, with clinical manifestations including skin and soft-tissue infections, sepsis, pneumonia and toxic shock syndrome. A large proportion of these infections can be attributed to globally disseminated, methicillin-resistant clones associated with healthcare settings (HA-MRSA), often arising in individuals with predisposing risk factors [1]. In the past two decades, community-associated methicillin-resistant *S. aureus* (CA-MRSA) clones distinct from HA-MRSA clones

have emerged to cause infections in otherwise healthy individuals and are often characterised by enhanced virulence and transmissibility [2]. Some of these clones have become a major cause of nosocomial infections, obscuring the distinction between CA-MRSA and HA-MRSA [3-5]. Although most research has so far originated from industrialised countries such as the United Kingdom and United States [2], there is an increasing awareness of the emergence of CA-MRSA clones in non- and newly-industrialised nations in Africa, Asia and the Indian sub-continent [6-11]. Some of these clones have now become globally disseminated [2,10].

Sequence type 772 (ST772) was originally isolated from India and Bangladesh, where it has become

* Correspondence: steven.tong@menzies.edu.au

¹Menzies School of Health Research, Darwin, Northern Territory, Australia
Full list of author information is available at the end of the article

increasingly common [7,9,12-14] and appears capable of invading healthcare environments and displacing previously successful nosocomial MRSA [7,15]. It has subsequently been identified in England, France, Germany, Ireland, Italy, Norway, Abu Dhabi, Saudi-Arabia, Malaysia, Australia and New Zealand [8], often in patients with family background or travel histories to India or Bangladesh [15-18]. In the past few years, several draft genomes have been assembled, revealing the genomic composition of ST772-MRSA-V [19-23]. This lineage is closely related to clonal complex CC1 and harbours the relatively small and mobile staphylococcal cassette chromosome *SCCmec* type V [22]. The draft genomes accommodate a novel prophage Φ -IND772PVL, carrying the enterotoxin gene *sea* and an operon encoding Pantone-Valentine Leukocidin (PVL), *lukS/F-PV* [21]. This potent combination of toxins on the same prophage has so far not been reported in other strains of *S. aureus* [21]. A heavily truncated *hlb*-converting phage is also present, which carries only *scn* (staphylococcal complement inhibitor) as part of the immune evasion cluster [21]. In addition, a variety of enterotoxins and superantigens have been detected in ST772-MRSA-V, including the variable presence of *sec* and *sel*, as well as the enterotoxin gene cluster *egc* [8,15,20]. Finally, several genes encoding antibiotic resistance determinants (e.g., against beta-lactams, aminoglycosides, fluoroquinolones, tetracyclines) and non-synonymous mutations in resistance-associated genes have been discovered, which correspond to previously identified multidrug resistant phenotypes of ST772-MRSA-V [7,15,20,24].

Despite these advances, the genomic location and context of resistance and virulence factors, which are frequently associated with mobile genetic elements (MGEs) [1,25], has not yet been unambiguously determined in ST772-MRSA-V [8,20]. Published draft genomes of ST772-MRSA-V have employed short-read sequencing technologies (i.e., Illumina), resulting in relatively fragmented assemblies (43–78 contigs, Table 1) and annotations and may lack important positional information of elements associated with resistance or virulence (e.g., resistance genes in [20]). Recently, third-generation sequencing technologies have allowed for the closure of complex genomic regions and the recovery of contiguous genome sequences [26,27], including for *S. aureus* [28-30].

In this study, we present the complete genome sequence of DAR4145, a multidrug resistant strain of ST772-MRSA-V from Mumbai, India [31]. Pacific Bioscience (PacBio) single-molecule real-time (SMRT) sequencing resulted in a single contig containing the complete chromosome of DAR4145, which was error checked and corrected with Illumina reads. Annotation and comparison to nine other isolates from Australia, Malaysia and India highlighted an association of resistance and virulence determinants with MGEs. Additional resistance-associated mutations and gene

redundancy explain the multidrug-resistant phenotype of DAR4145. The closed genome sequence fully defines the genomic composition of DAR4145 and provides a valuable reference genome for future investigations into the genomic epidemiology and phylogeography of ST772-MRSA-V.

Results and discussion

SMRT sequencing and subsequent assembly generated a single continuous sequence containing the complete chromosome of DAR4145 (Figure 1). A second shorter sequence was identical to the plasmid pKH-18, carrying cadmium resistance gene *cadD* and its regulator *cadX* [GenBank: EU333812.1]. The length of the completed chromosome was 2,860,508 base pairs (bp) with a G + C content of 32.85%. 2,642 protein coding regions, as well as 18 rRNAs, 61 tRNAs, 25 ncRNAs (including *RNAIII* and *sprD*) and 35 transposases associated with insertion sequence elements and transposons (including partial and putative sequences) were identified and annotated. Notably, gaps in the draft genome assemblies of other strains of ST772-MRSA-V (Figure 1) frequently occurred in the vicinity of transposases, suggesting that some of the short read assemblies were unable to bridge gaps associated with repetitive genomic elements. This clearly emphasizes the potential for obtaining fully defined *de novo* assemblies using third-generation sequencing technologies, which are able to resolve both local and global repeats [32] and avoid the introduction of reference bias. *In silico* multi-locus sequencing typing confirmed ST772, with a single allele divergence from ST1 (*pta-22*). Although a single locus variant of ST1 (and thus part of clonal complex 1), the *agr* locus of DAR4145 is part of *agr* group II rather than group III. A close relationship of the core genome is evident for all available genomes of ST772-MRSA-V. Excluding MGEs, the pairwise core genome SNP differences between DAR4145 and other ST772-MRSA-V genomes ranged from 52 to 254 SNPs (Table 1). DAR4145 harbours the staphylococcal cassette chromosome *SCCmec* type V containing the cassette recombinase *ccrC2*; the composition and arrangement are nearly identical to the previously identified *SCCmec* in strain 118 [22], with the exception of an additional mobile element inserted immediately upstream of the terminal direct repeat DR2. The element carried the bifunctional aminoglycoside modifying gene *aacA-aphD* and was flanked by transposases of IS256. Its composition identified the element as transposon Tn4001 [33] [GenBank: AB682805.1] and it was present at the same location in all available genomes of ST772-MRSA-V, except strain 118 (Figure 1).

Resistance to multiple antibiotics mediated by mobile genetic elements

We found several additional resistance-associated genes, most of which were located on MGEs and correlated

Table 1 Draft genomes of ST772-MRSA-V used for comparison with DAR4145

Strain	Origin	Year of isolation	Core genome SNPs vs. DAR4145	Number of contigs	Publication	Genbank accession
07-01748	Western Australia	2007	72	70	Monecke et al. [20]	AZBT000000000
KT/Y21	Terrenganu, Malaysia	2009	73	70	Suhaili et al. [23]	AOCQ000000000
118	Bangalore, India	2008	70	78	Prabhakara et al. [21]	AJGE000000000
120	Bangalore, India	2009	52	56	Prabhakara et al. [21]	ALWE000000000
333	Madurai, India	2010	78	54	Prabhakara et al. [21]	ALWF000000000
VH60	Bangalore, India	2007	65	44	Prabhakara et al. [21]	ALWG000000000
3989	Hyderabad, India	2007	79	43	Prabhakara et al. [21]	ALWH000000000
LVP2	Bhubaneshwar, India	2010	119	73	Balakuntla et al. [22]	AOFV000000000
3957	Hyderabad, India	2007	254	43	Balakuntla et al. [22]	AOFU000000000

with the resistance phenotype of DAR4145 (Table 2). One core genome copy of the dihydrofolate reductase gene *dhfrA* and two copies of *dhfrG* were identified, which have been linked to high-level trimethoprim resistance in *S. aureus* [34]. The first copy of *dhfrG* was found in the context of several phage genes and within the boundary of the terminal attachment site (*attR*) of a putative SaPI, partially identical with the pathogenicity island SaPITokyo12381 (BLASTn 99% identity, 70% coverage) [GenBank: AB860418.1]. The island also contained the enterotoxin genes *sec* and *sel*, placed in a similar context and position as ν Sa3 Type II, a pathogenicity island of strain MW2 [35]. Monecke et al. [8] noted the frequent but variable presence of *sec* and *sel* in ST772-MRSA-V, which was supported in this study by the absence of the putative SaPI in strain 3989 (Figure 1). The second copy of *dhfrG* was found in the chromosome at 2,474,700 bp. Both sequences, with two adjacent coding regions for hypothetical proteins, were identical to genomic regions in ST239 strain TW20 from the UK [36], the trimethoprim-resistant *S. pyogenes* strain A1085 from India [37] and part of the transposon Tn6198 from *Listeria monocytogenes* [38] [GenBank: FN433596.1, GenBank: JX498941.1, GenBank: JX120102.1]. Analysis with ResFinder [39] confirmed the presence of the two *dhfrG* copies in DAR4145, but only identified the SaPI-associated copy in the draft genomes of 07–17048, 118, 3957 and KT/Y21 and the core-associated copy in isolates 120, 333, 3989, LVP2, VH60. The presence of *dhfrG* was also confirmed by mapping paired-end short reads against our resistome database and corresponded to a high-level trimethoprim/sulfamethoxazole (TMP-SXT) resistance phenotype of DAR4145 (MIC \geq 320 μ g/L) (Table 2). The *dhfrG* gene was previously thought to be rare in *S. aureus* colonising humans, but has recently been found to be widespread in sub-Saharan Africa [34].

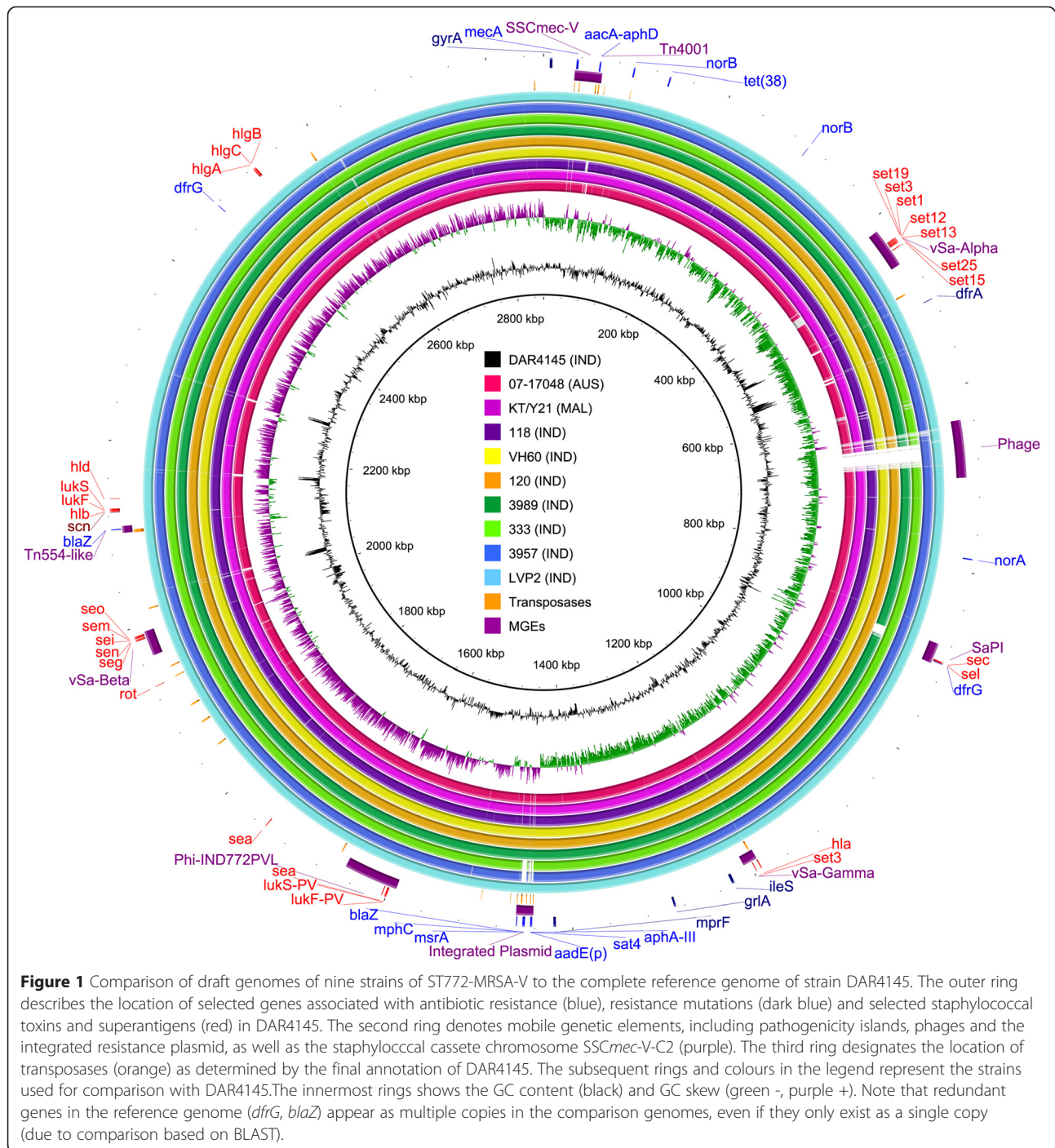
Similarly, two copies of the beta-lactam resistance operon (*blaZ*, *blaI*, *blaR*) were identified in DAR4145. The first

copy was part of a partial integrated plasmid highly similar in composition to the plasmid p18810-03 (BLASTn 99% identity, 85% coverage) [GenBank: CP002141.1] found in CA-MRSA strain 18810 of USA300 [40] and nearly identical with the integrated plasmid in strain 11819–97 of CA-MRSA ST80-IV (BLASTn 99% identity, 99% coverage) [GenBank: CP003194.1] [28]. The integrated plasmid encoded a resistance cluster composed of *aphA-III*, *sat4* and a partial *aadE* (aminoglycoside resistance), as well as *mphC* and *msrA* (macrolide resistance), thus resolving the genomic context of these genes previously noted by Monecke et al. [20]. Draft genomes of strains 07–17048 and 120 respectively lacked or carried a truncated, plasmid-associated *blaZ*, whereas the integrated plasmid was entirely absent in strains 333 and 3957 (Figure 1, Table 2). The second copy of the beta-lactam resistance operon was located immediately downstream of three transposases related to Tn554 and was present in all isolates of ST772-MRSA-V. The Tn554-like transposon [41] was previously found with nearly identical sequence composition in H-EMRSA-15 [CP007659.1] and HO 5096 0412 [GenBank: HE681097.1] of ST22 [41,42].

Resistance to multiple antibiotics mediated by core chromosome genes and mutations

Three additional, chromosomal resistance genes identified in DAR4145 were the transporter genes *norA* and *norB*, whose altered expression is associated with quinolone resistance [43,44] and the efflux pump *tet(38)*, which can confer resistance to tetracyclines when upregulated [43,45,46]. The presence of *tet(38)* alone did not result in phenotypic resistance for DAR4145 (Table 2). Notably, *tet(38)* also facilitates the efflux of antibacterial fatty acids, which promotes survival in an abscess environment and colonisation of skin surfaces [46].

Furthermore, the chromosomal genes *grrA/grrB* and *gyrA/gyrB* harbour common non-synonymous mutations associated with increased resistance against quinolones



in *S. aureus* [47-49]. We found two mutations in ST772-MRSA-V, one in *grlA* (S80Y, S80F in isolates 3957 and LVP2) and another in *gyrA* (S84L, in all except isolate 118). These correspond to ciprofloxacin-resistance associated mutations previously found in ST772-MRSA-V and other STs from India [24]. We also detected non-synonymous mutations in the genes *dfrA* (V134I), *ileS-1* (D621H) and *mprF* (L335S) that were present in all examined strains of ST772-MRSA-

V. These were located in the vicinity of mutations that have previously been linked with resistance to trimethoprim [50], mupirocin [51] and daptomycin [34] in *S. aureus*, respectively.

Phage and virulence determinants

In addition to the antibiotic resistance factors and the putative SaPI, DAR4145 harbours the genomic islands *vSa α* , *vSa β* and *vSa γ* . These were present in all investigated

Table 2 Selected phenotypic and genomic antibiotic resistance profiles in strains of ST772-MRSA-V

Antibiotic	Phenotype (MIC µg/L)		Resistance Genes and Mutations in ST772-MRSA-V									
	DAR4145		DAR4145	07-17048	KT/Y21	118	120	333	3957	3989	LVP2	VH60
Oxacillin	R (≥4)	<i>mecA</i> (M)	+	+	+	+	+	+	+	+	+	+
		<i>blaZ</i> (M)	+	+	+	+	+	+	+	+	+	+
		<i>blaZ</i> (P)	+	-	+	+	+/-	-	-	+	+	+
Gentamicin	R (≥16)	<i>aacA-aphD</i> (M)	+	+	+	-	+	+	+	+	+	+
		<i>aphA-III, sat4, aadE(p)</i> (P)	+	+	+	+/-	+	-	-	+	+	+
Trimethoprim/Sulfamethoxazole	R (≥320)	<i>dfrA</i> (C)	+	+	+	+	+	+	+	+	+	+
		<i>dfrG</i> (C)	+	-	-	-	+	+	-	+	+	+
		<i>dfrG</i> (M)	+	+	+	+	-	-	+	-	-	-
Ciprofloxacin Levofloxacin Moxifloxacin	R (4) R (4) I (1)	<i>norA/norB</i> (C)	+	+	+	+	+	+	+	+	+	+
		<i>grlA</i> (S80T/F)	+	+	+	+	+	+	+	+	+	+
		<i>gyrA</i> (S84L)	+	+	+	+	+	+	+	+	+	+
Tetracycline	S (≤1)	<i>tet(38)</i> (C)	+	+	+	+	+	+	+	+	+	
Erythromycin Clindamycin	I (1) S (≤0.25)	<i>msrA-mphC</i> (P)	+	+	+	+	+	-	-	+	+	+

(R = Resistant, I = Intermediate, S = Susceptible, C = Chromosomal, M = Mobile Genetic Element, P = Integrated Resistance Plasmid, + = present, - = absent, p = partial, +/- = partial or truncated).

genomes of ST772-MRSA-V (Figure 1) and encoded several superantigen-like proteins and toxin determinants, including the enterotoxin gene cluster (*egc*) on ν Sa β as well as α -haemolysin (*hla*) and exfoliative toxin A (*eta*) on ν Sa γ . In addition, ν Sa α harbors *aadK*, encoding aminoglycoside 6-adenylyltransferase, which is associated with low-level streptomycin resistance in *Bacillus subtilis* [52]. The recently described prophage Φ -IND772PVL, containing the staphylococcal enterotoxin A (*sea*) and the PVL operon (*lukF-PV*, *lukS-PV*), was also found in all isolates, confirming its widespread presence in ST772-MRSA-V from India [21]. DAR4145 harboured the heavily truncated *hly*-converting prophage encoding the staphylococcal complement inhibitor (*scn*) [20,21], immediately followed by the *hlg*-related leukocidin genes *lukF* and *lukS* [53]. We also found a second putative enterotoxin gene with high similarity to *sea* within the chromosome at 1,746,983 bp (BLASTn, 98% identity, 98% coverage in CA-MRSA ST80-IV strain 11819–97) [GenBank CP003194.1]. Additionally, DAR4145 contained one previously undetected prophage region at 637,316 bp. This prophage did not harbour known resistance or virulence genes and the central region was present only in DAR4145, 3957 and LVP2 (Figure 1). The entire region was similar to ϕ Mu50B of strain Mu50 (BLASTn, 95% identity, 83% of total coverage) [GenBank: BA000017.4], with the core region showing highest similarity to ϕ NM1 and ϕ NM2 (BLASTn, 96% identity, 51–54% of total coverage) [GenBank: DQ530359.1 and DQ530360.1].

Finally, a putative and a truncated transposase for IS1272 were found in the vicinity (–477 bp) of the global transcriptional regulator of virulence *rot* (repressor of toxins). Benson et al. [54] demonstrated that insertion of IS257 elements into the promoter region of *rot* was responsible for derepression of cytotoxin expression and increased virulence of USA500 strains. Although we did not examine the functional significance of this arrangement, the role of gene expression regulation is becoming increasingly recognized to contribute to variations in virulence and antibiotic resistance of *S. aureus* [43,46,54,55] and thus may be of interest to future investigations into ST772-MRSA-V. Similarly, the functional role of the *agr* group II locus in DAR4145 is of interest as it differs from the *agr* group III locus found in other clonal complex 1 strains [56].

Comparison to other STs

There are notable differences between DAR4145 and other sequenced CA-MRSA strains with regards to acquired antimicrobial resistance and virulence determinants (Additional file 1). The level of multi-drug resistance is greater than that seen in the key CA-MRSA clones from the US – USA300 [57] and MW2 [35], Europe – ST80 [28], Australia – ST93 [58], and Asia – ST59 [59]. In particular, DAR4145

has MGEs encoding for resistance to cotrimoxazole and aminoglycosides that are not present in the other CA-MRSA genomes. The integrated plasmid harbouring resistance genes *blaZ*, *mphC*, *aphA-III* and *msrA* is found only in DAR4145 and ST80 strain 11819–97; although in contrast to 11819–97 the integrated plasmid is not within the *SCCmec* element in DAR4145 [28].

Although all the above listed CA-MRSA strains harbor PVL-encoding phage, the combination of *sea* and PVL on Φ -IND772PVL only occurs in ST772-MRSA-V [21]. In comparison to the other strains that carry 2–7 enterotoxin genes, DAR4145 harbors 10 such genes, including *sec* and *sel* on the putative SaPI and the enterotoxin cluster *seg*, *sei*, *sem*, *sen* and *seo* on ν Sa β . DAR4145 contains capsid type 5, but lacks both the serine-like protease operon *spl* and *lukD/E*. In total, compared to other important genome sequenced CA-MRSA strains, DAR4145 contains a striking combination of acquired antimicrobial and virulence determinants. It should be noted that the small number of available genomes examined in this study may cause a bias towards isolates that are particularly virulent or multi-drug resistant and therefore may not be an accurate assessment of these traits in the lineage at large. However, previous studies on larger collections that included molecular typing of resistance and virulence factors have determined similar resistance profiles and confirmed the widespread presence of virulence determinants, such as the *egc* gene cluster or PVL [7,15,21].

Conclusions

The complete genome of ST772-MRSA-V is remarkable for the presence of multiple MGEs mediating a highly antibiotic resistant phenotype. This phenotype is further augmented by the presence of core genome resistance-associated mutations and redundancy of some resistance mediating genes. For example, DAR4145 demonstrates high-level resistance to TMP-SXT (MIC \geq 320 μ g/L) and contains three relevant resistance genes, *dfrA* and two copies of *dfrG*. Additionally, DAR4145 carries a large complement of virulence factors including multiple enterotoxins, α -haemolysin and PVL.

Recommended oral antibiotics for the treatment of community-associated *S. aureus* skin and soft tissue infections include anti-staphylococcal β -lactams, clindamycin, TMP-SXT, erythromycin, and doxycycline [60]. DAR4145 is resistant to all of these antibiotics except for doxycycline (and even then it carries *tet(38)*, which if up-regulated could conceivably result in phenotypic resistance). Unfortunately, a multidrug-resistant genotype is not restricted to DAR4145 and appears to be a consistent feature of all ST772-MRSA-V strains analysed in this study. Thus, simple oral treatment options for the ST772 lineage are truly limited.

The increasing prevalence of ST772-MRSA-V in India and its appearance in multiple regions of the world [7-9,12-14,20], should raise concerns. As with highly resistant Gram-negatives, drivers of the emergence and spread of ST772-MRSA-V need urgent study. Here, we have defined the bacterial factors that contribute to the multidrug-resistant phenotype of this lineage. Studies to understand the global phylogeography of ST772 can now make use of the closed reference genome of DAR4145.

Methods

Sequencing and assembly

Strain DAR4145 was recovered as a blood culture isolate in a patient from Mumbai, India in 2009 [31]. This strain was multilocus sequence type 772, *spa* type 1080 (Ridom *spa* type t657), PVL+, SCC*mec* V, and *dru* type 10bk. For this study, antimicrobial susceptibility testing was performed with the Vitek2 platform using 22359 VITEK AST-P612 cards (bioMérieux, France) with Clinical and Laboratory and Standards Institute breakpoints utilised [61].

Sequencing for *de novo* assembly was performed on the PacBio RS (Menlo Park, CA, USA). A sequencing library (SMRTbell) was generated from approximately 2 µg of genomic DNA sheared to an average fragment length of 10 kb. Library preparation was carried out following manufacturer protocols using DNA Template Prep Kit 2.0 (3 – 10 Kb). The sequencing enzyme used was version 2.0 from the DNA/Polymerase Binding Kit 2.0, movie lengths were 90 min, and sequencing chemistry was version C2. From eight SMRT cells of sequencing, we generated 313,192 polymerase reads with an average length of 3.1 kb (and a total of 401,221 sub-reads with an average length of 2.0 kb, the largest was 14.7 kb; sub-reads are the individual forward and reverse reads of every SMRTbell library molecule excised from a polymerase read). *De novo* assembly was performed with HGAP v.2.2. Circularisation was achieved by manual comparison and removal of a region of overlap. For short-read re-mapping and error-correction, an Illumina library with a median insert size of 238 bp (first and third quartile: 194/303 bp) was generated and sequenced on the Illumina HiSeq 2000 with 100 bp paired-end runs, producing 3,826,612 sequencing reads.

Computational analysis

Automated error correction was performed by mapping paired-end short reads against the PacBio assembly with five iterations in iCORN v.2 [62]. Results were confirmed with SPANDEX v.2.4, using standard parameters [63]. The corrected reference sequence was annotated with Prokka v1.07 [64] and supplemented with results from the annotation servers BASys [65] and RAST [66]. Visual inspection and manual curation was carried

out with Artemis v.16.0.0 [67,68] and BLAST searches against databases from NCBI [69]. Virulence factors and phages were identified using VirulenceFinder 1.2 (90% identity) [70] and the PHAST server [71]. *In silico* multi-locus sequence typing (MLST) was performed using the MLST server v1.3 [72] and validated by mapping paired-end short reads to a previously assembled MLST database from Sanger Institute.

We obtained publicly available assemblies and short reads for nine additional whole-genome sequences of ST772-MRSA-V (07-17048, 118, 120, 333, 3957, 3989, VH60, LVP2, KT/Y21; see Table 1) for comparison and visualization with the BLAST Ring Image Generator (BRIGS) v.0.95 [73]. Each genome was also mapped against DAR4145 using SMALT v.0.7.5 with default settings, followed by exclusion of regions of known MGEs. High-quality core genome SNPs were identified using a combination of SAMtools v.1.1 mpileup and bcftools, with filters applied as previously described [74]. Genes and high-quality SNPs implicated with antimicrobial resistance were identified in all strains by mapping available short read data against a database of known, annotated sequences from Sanger Institute with SMALT v.0.7.5 (Sanger Institute) and by using an in-house pipeline combining GATK v.3.3.0 [75,76] and SAMtools v.1.1 [77,78]. Resistance genes were verified with ResFinder v.2.1 [39] at 80% minimum length and 90% identity cut-offs, using assemblies ordered against the reference DAR4145 in Mauve v.2.3.1 [79] in order to determine their location and genomic context. Specific regions of interest (i.e., resistance genes, phages, virulence factors) were further investigated and validated by mapping available short reads of isolates, including DAR4145, against the PacBio assembly with standard parameters in SPANDEX. Finally, we obtained representative genomes of five significant STs for comparison with ST772-MRSA-V in ResFinder and Virulence Finder: MW2 [GenBank: BA000033.2], USA300 FPR3757 [GenBank: CP000255.1], ST59 M013 [GenBank: CP0003166.1], ST80 11819-97 [GenBank: CP003194.1] and ST93 JKD6159 [GenBank: CP002114.2]. Results were visualized as a presence/absence matrix in MS Excel 2010.

Ethics statement

The bacterial isolate DAR4145 was collected as part of routine clinical service provision [31]. The use of a de-identified bacterial strain described in this manuscript, and the study of the bacterial isolates and not human subjects, meant that formal Human Ethics Committee approval or Informed Patient Consent was not required.

Availability of supporting data

The complete genome assembly and annotation of DAR4145 has been submitted to GenBank under the

accession number CP010526. Illumina reads are available at the European Nucleotide Archive under the accession number ERS161279.

Additional file

Additional file 1: Comparison of resistance and virulence genes in representative genomes of five significant STs and ST772-MRSA-V DAR4145.

Abbreviations

MGE: Mobile genetic element; MLST/ST: Multilocus sequence type/sequence type; (CA-/HA-) MRSA: (community-associated-/healthcare-associated) methicillin-resistant *Staphylococcus aureus*; PacBio: Pacific biosciences; SaPI: *Staphylococcus aureus* Pathogenicity Island; SMRT: Single-molecule real-time sequencing; TMP-SXT: Trimethoprim/sulfamethoxazole.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

EJS carried out the computational analysis and drafted the manuscript, PA participated in the analysis and helped to draft the manuscript, DSS participated in the analysis, SRH, PC, MTGH, JP and SDB carried out the sequencing and genome assembly, AM and DAR provided the sample and information on strain type and antimicrobial susceptibility and participated in the analysis, SYCT conceived of the study, performed analysis, designed and coordinated the study and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

SYCT is an Australian National Health and Medical Research Council Career Development Fellow (1065736). DAR was supported in part by National Institutes of Health grant GM080602. SRH, PC, MTGH, JP and SDB were supported by Wellcome Trust grant 098051.

Author details

¹Menzies School of Health Research, Darwin, Northern Territory, Australia. ²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. ³Pushpagiri Research Center, Pushpagiri Institute of Medical Sciences and Research Center, Thiruvalla, India. ⁴School of Medicine, University of St. Andrews, St. Andrews, UK. ⁵Department of Microbiology, University of Mississippi Medical Center, Jackson, MS, USA.

Received: 22 January 2015 Accepted: 28 April 2015

Published online: 16 May 2015

References

- Stefani S, Chung DR, Lindsay JA, Friedrich AW, Kearns AM, Westh H, et al. Methicillin-resistant *Staphylococcus aureus* (MRSA): global epidemiology and harmonisation of typing methods. *Int J Antimicrob Agents*. 2012;39:273–82.
- Mediavilla JR, Chen L, Mathema B, Kreiswirth BN. Global epidemiology of community-associated methicillin resistant *Staphylococcus aureus* (CA-MRSA). *Curr Opin Microbiol*. 2012;15:588–95.
- Seybold U, Kourbatova EV, Johnson JG, Halvosa SJ, Wang YF, King MD, et al. Emergence of community-associated methicillin-resistant *Staphylococcus aureus* USA300 genotype as a major cause of health care-associated blood stream infections. *Clin Infect Dis*. 2006;42:647–56.
- Popovich KJ, Weinstein RA, Hota B. Are community-associated methicillin-resistant *Staphylococcus aureus* (MRSA) strains replacing traditional nosocomial MRSA strains? *Clin Infect Dis*. 2008;46:787–94.
- Otter JA, French GL. Community-associated methicillin-resistant *Staphylococcus aureus* strains as a cause of healthcare-associated infection. *J Hosp Infect*. 2011;79:189–93.
- Ghebremedhin B, Olugbosi MO, Raji AM, Layer F, Bakare RA, König B, et al. Emergence of a community-associated methicillin-resistant *Staphylococcus aureus* strain with a unique resistance profile in Southwest Nigeria. *J Clin Microbiol*. 2009;47:2975–80.
- D'Souza N, Rodrigues C, Mehta A. Molecular characterization of methicillin-resistant *Staphylococcus aureus* with emergence of epidemic clones of sequence type (ST) 22 and ST 772 in Mumbai, India. *J Clin Microbiol*. 2010;48:1806–11.
- Monecke S, Coombs G, Shore AC, Coleman DC, Akpaka P, Borg M, et al. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. *PLoS One*. 2011;6, e17936.
- Shambat S, Nadig S, Prabhakara S, Bes M, Etienne J, Arakere G. Clonal complexes and virulence factors of *Staphylococcus aureus* from several cities in India. *BMC Microbiol*. 2012;12:64.
- Tong SYC, Kearns AM. Community-associated MRSA from the Indian subcontinent. *Lancet Infect Dis*. 2013;13:734–5.
- Falagas ME, Karageorgopoulos DE, Leptidis J, Korbila IP. MRSA in Africa: filling the global map of antimicrobial resistance. *PLoS One*. 2013;8, e68024.
- Goering RV, Shawar RM, Scangarella NE, O'Hara FP, Amrine-Madsen H, West JM, et al. Molecular epidemiology of methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* isolates from global clinical trials. *J Clin Microbiol*. 2008;46:2842–7.
- Afroz S, Kobayashi N, Nagashima S, Alam MM, Hossain ABMB, Rahman MA, et al. Genetic Characterization of *Staphylococcus aureus* Isolates Carrying Panton-Valentine Leukocidin Genes in Bangladesh. *Jpn J Infect Dis*. 2008;61:393–6.
- Nadig S, Velusamy N, Lalitha P, Kar S, Sharma S, Arakere G. *Staphylococcus aureus* eye infections in two Indian hospitals: emergence of ST772 as a major clone. *Clin Ophthalmol*. 2012;6:165–73.
- Brennan GI, Shore AC, Corcoran S, Tecklenborg S, Coleman DC, O'Connell B. Emergence of hospital- and community-associated panton-valentine leukocidin-positive methicillin-resistant *Staphylococcus aureus* genotype ST772-MRSA-V in Ireland and detailed investigation of an ST772-MRSA-V cluster in a neonatal intensive care unit. *J Clin Microbiol*. 2012;50:841–7.
- Ellington MJ, Ganner M, Warner M, Cookson BD, Kearns AM. Polyclonal multiply antibiotic-resistant methicillin-resistant *Staphylococcus aureus* with Panton-Valentine leukocidin in England. *J Antimicrob Chemother*. 2010;65:46–50.
- Coombs GW, Monecke S, Pearson JC, Tan H, Chew Y-K, Wilson L, et al. Evolution and diversity of community-associated methicillin-resistant *Staphylococcus aureus* in a geographical region. *BMC Microbiol*. 2011;11:215.
- Zanger P, Nurjadi D, Schleucher R, Scherbaum H, Wolz C, Kreamer PG, et al. Import and spread of Panton-Valentine Leukocidin-positive *Staphylococcus aureus* through nasal carriage and skin infections in travelers returning from the tropics and subtropics. *Clin Infect Dis*. 2012;54:483–92.
- Prabhakara S, Khedkar S, Loganathan RM, Chandana S, Gowda M, Arakere G, et al. Draft genome sequence of *Staphylococcus aureus* 118 (ST772), a major disease clone from India. *J Bacteriol*. 2012;194:3727–8.
- Monecke S, Baier V, Coombs GW, Slickers P, Ziegler A, Ehrlich R. Genome sequencing and molecular characterisation of *Staphylococcus aureus* ST772-MRSA-V, "Bengal Bay Clone". *BMC Res Notes*. 2013;6:548.
- Prabhakara S, Khedkar S, Shambat SM, Srinivasan R, Basu A, Norrby-Teglund A, et al. Genome sequencing unveils a novel sea enterotoxin-carrying PVL phage in *Staphylococcus aureus* ST772 from India. *PLoS One*. 2013;8, e60013.
- Balakuntla J, Prabhakara S, Arakere G. Novel rearrangements in the staphylococcal cassette chromosome mec type V elements of Indian ST772 and ST672 methicillin resistant *Staphylococcus aureus* strains. *PLoS One*. 2014;9, e94293.
- Suhaili Z, Lean S-S, Yahya A, Mohd Desa MN, Ali AM, Yeo CC. Draft Genome Sequence of Methicillin-Resistant *Staphylococcus aureus* KT/Y21, a Sequence Type 772 (ST772) Strain Isolated from a Pediatric Blood Sample in Terengganu, Malaysia. *Genome Announc*. 2014;2:e00271–14.
- Chakrakodi B, Prabhakara S, Nagaraj S, Etienne J, Arakere G. High Prevalence of Ciprofloxacin Resistance in Community Associated *Staphylococcus aureus* in a Tertiary Care Indian Hospital. *Adv Microbiol*. 2014;04:133–41.
- Lindsay JA, Holden MTG. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol*. 2004;12:378–85.
- Huddleston J, Ranade S, Malig M, Antonacci F, Chaisson M, Hon L, et al. Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res*. 2014;24:688–96.
- Limmathurotsakul D, Holden MTG, Coupland P, Price EP, Chantratita N, Wuthiekanun V, et al. Microevolution of *Burkholderia pseudomallei* during an acute infection. *J Clin Microbiol*. 2014;52:3418–21.
- Stegger M, Price LB, Larsen AR, Gillette JD, Waters AE, Skov R, et al. Genome sequence of *Staphylococcus aureus* strain 11819–97, an ST80-IV European community-acquired methicillin-resistant isolate. *J Bacteriol*. 2012;194:1625–6.

29. Altman DR, Sebra R, Hand J, Attie O, Deikus G, Carpini KWD, et al. Transmission of Methicillin-Resistant *Staphylococcus aureus* via Deceased Donor Liver Transplantation Confirmed by Whole Genome Sequencing. *Am J Transplant*. 2014;14:2640–4.
30. Parker D, Narechania A, Sebra R, Deikus G, Larussa S, Ryan C, et al. Genome Sequence of Bacterial Interference Strain *Staphylococcus aureus* 502A. *Genome Announc*. 2014;2:e00284–14.
31. Manoharan A, Zhang L, Poojary A, Bhandarkar L, Koppikar G, Robinson DA. An outbreak of post-partum breast abscesses in Mumbai, India caused by ST22-MRSA-IV: genetic characteristics and epidemiological implications. *Epidemiol Infect*. 2012;140:1809–12.
32. Koren S, Phillippy AM. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr Opin Microbiol*. 2015;23:110–20.
33. Lyon BR, May JW, Skurray RA. Tn4001: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Mol Gen Genet*. 1984;193:554–6.
34. Nurjadi D, Olalekan AO, Layer F, Shittu AO, Alabi A, Ghebremedhin B, et al. Emergence of trimethoprim resistance gene *dfgG* in *Staphylococcus aureus* causing human infection and colonization in sub-Saharan Africa and its import to Europe. *J Antimicrob Chemother*. 2014;69:2361–8.
35. Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, et al. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet*. 2002;359:1819–27.
36. Holden MTG, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, 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;192:888–92.
37. Bergmann R, Sagar V, Nitsche-Schmitz DP, Chhatwal GS. First detection of trimethoprim resistance determinant *dfgG* in *Streptococcus pyogenes* clinical isolates in India. *Antimicrob Agents Chemother*. 2012;56:5424–5.
38. Bertsch D, Uruty A, Anderegg J, Lacroix C, Perreten V, Meile L. Tn6198, a novel transposon containing the trimethoprim resistance gene *dfgG* embedded into a Tn916 element in *Listeria monocytogenes*. *J Antimicrob Chemother*. 2013;68:986–91.
39. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012;67:2640–4.
40. Kennedy AD, Porcella SF, Martens C, Whitney AR, Braughton KR, Chen L, et al. Complete nucleotide sequence analysis of plasmids in strains of *Staphylococcus aureus* clone USA300 reveals a high level of identity among isolates with closely related core genome sequences. *J Clin Microbiol*. 2010;48:4504–11.
41. Holden MTG, Hsu L-Y, Kurt K, Weinert LA, Mather AE, Harris SR, et al. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res*. 2013;23:653–64.
42. Sabirova JS, Xavier BB, Hernalsteens J-P, De Greve H, Ieven M, Goossens H, et al. Complete Genome Sequences of Two Prolific Biofilm-Forming *Staphylococcus aureus* Isolates Belonging to USA300 and EMRSA-15 Clonal Lineages. *Genome Announc*. 2014;2:e00610–4.
43. Truong-Bolduc QC, Dunman PM, Strahilevitz J, Projan SJ, Hooper DC. MgrA is a multiple regulator of two new efflux pumps in *Staphylococcus aureus*. *J Bacteriol*. 2005;187:2395–405.
44. Kwak YG, Truong-Bolduc QC, Bin Kim H, Song K-H, Kim ES, Hooper DC. Association of *norB* overexpression and fluoroquinolone resistance in clinical isolates of *Staphylococcus aureus* from Korea. *J Antimicrob Chemother*. 2013;68:2766–72.
45. Ding Y, Onodera Y, Lee JC, Hooper DC. *NorB*, an efflux pump in *Staphylococcus aureus* strain MW2, contributes to bacterial fitness in abscesses. *J Bacteriol*. 2008;190:7123–9.
46. Truong-Bolduc QC, Villet RA, Estabrooks ZA, Hooper DC. Native efflux pumps contribute resistance to antimicrobials of skin and the ability of *Staphylococcus aureus* to colonize skin. *J Infect Dis*. 2014;209:1485–93.
47. Ferrero L, Cameron B, Crouzet J. Analysis of *gyrA* and *griA* mutations in stepwise-selected ciprofloxacin-resistant mutants of *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 1995;39:1554–8.
48. Yamagishi J, Kojima T, Oyamada Y, Fujimoto K, Hattori H, Nakamura S, et al. Alterations in the DNA topoisomerase IV *griA* gene responsible for quinolone resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 1996;40:1157–63.
49. Schmitz FJ, Jones ME, Hofmann B, Hansen B, Scheuring S, Lückefahr M, et al. Characterization of *griA*, *griB*, *gyrA*, and *gyrB* mutations in 116 unrelated isolates of *Staphylococcus aureus* and effects of mutations on ciprofloxacin MIC. *Antimicrob Agents Chemother*. 1998;42:1249–52.
50. Vickers AA, Potter NJ, Fishwick CWG, Chopra I, O'Neill AJ. Analysis of mutational resistance to trimethoprim in *Staphylococcus aureus* by genetic and structural modelling techniques. *J Antimicrob Chemother*. 2009;63:1112–7.
51. Antonio M, McFerran N, Pallen MJ. Mutations Affecting the Rossman Fold of Isoleucyl-tRNA Synthetase Are Correlated with Low-Level Mupirocin Resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2002;46:438–42.
52. Noguchi N, Sasatsu M, Kono M. Genetic mapping in *Bacillus subtilis* 168 of the *aadK* gene which encodes aminoglycoside 6-adenylyltransferase. *FEMS Microbiol Lett*. 1993;114:47–52.
53. Kaneko J, Kamio Y. Bacterial two-component and hetero-heptameric pore-forming cytolytic toxins: structures, pore-forming mechanism, and organization of the genes. *Biosci Biotechnol Biochem*. 2004;68:981–1003.
54. Benson MA, Ohneck EA, Ryan C, Alonzo F, Smith H, Narechania A, et al. Evolution of hypervirulence by a MRSA clone through acquisition of a transposable element. *Mol Microbiol*. 2014;93:664–81.
55. Kaito C, Saito Y, Ikuo M, Omae Y, Mao H, Nagano G, et al. Mobile genetic element SCCmec-encoded *psm-mec* RNA suppresses translation of *agrA* and attenuates MRSA virulence. *PLoS Pathog*. 2013;9, e1003269.
56. Robinson DA, Monk AB, Cooper JE, Feil EJ, Enright MC. Evolutionary Genetics of the Accessory Gene Regulator (*agr*) Locus in *Staphylococcus aureus*. *J Bacteriol*. 2005;187(24):8312–21.
57. Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, et al. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet*. 2006;367:731–9.
58. Chua K, Seemann T, Harrison PF, Davies JK, Coutts SJ, Chen H, et al. Complete genome sequence of *Staphylococcus aureus* strain JKD6159, a unique Australian clone of ST93-IV community methicillin-resistant *Staphylococcus aureus*. *J Bacteriol*. 2010;192:5556–7.
59. Huang T-W, Chen F-J, Miu W-C, Liao T-L, Lin A-C, Huang I-W, et al. Complete genome sequence of *Staphylococcus aureus* M013, a *pvl*-positive, ST59-SCCmec type V strain isolated in Taiwan. *J Bacteriol*. 2012;194:1256–7.
60. Stevens DL, Bisno AL, Chambers HF, Dellinger EP, Goldstein EJC, Gorbach SL, et al. Practice guidelines for the diagnosis and management of skin and soft tissue infections: 2014 update by the Infectious Diseases Society of America. *Clin Infect Dis*. 2014;59:e10–52.
61. CLSI. Performance standards for antimicrobial susceptibility testing; twenty-first informational supplement M100-S21. Wayne, PA: CLSI M100-S21; 2011. p. 1–172.
62. Otto TD, Sanders M, Berriman M, Newbold C. Iterative Correction of Reference Nucleotides (iCORN) using second generation sequencing technology. *Bioinformatics*. 2010;26(14):1704–7.
63. Sarovich D, Price E. SPANX: a genomics pipeline for comparative analysis of large haploid whole genome re-sequencing datasets. *BMC Res Notes*. 2014;7:618.
64. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9.
65. Van Domselaar GH, Stothard P, Shrivastava S, Cruz JA, Guo A, Dong X, et al. BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res*. 2005;33(2):W455–9.
66. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. 2008;9:386.
67. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics*. 2000;16:944–5.
68. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*. 2012;28(4):464–9.
69. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
70. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic *Escherichia coli*. *J Clin Microbiol*. 2014;52(5):1501–10.
71. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A Fast Phage Search Tool. *Nucleic Acids Res*. 2011;39:W347–52.

72. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol.* 2012;50:1355–61.
73. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011;12:402.
74. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, et al. Evolution of MRSA during hospital transmission and intercontinental spread. *Science.* 2010;327:469–74.
75. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
76. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491–8.
77. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
78. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics.* 2011;27:2987–93.
79. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 2004;14:1394–403.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

