

# Genomic Portrait of Community-Associated Methicillin-Resistant *Staphylococcus Aureus* ST772-SCC*mec* V Lineage From India

**Yamuna Bakthavatchalam**

Christian Medical College & Hospital

**Karthick Vasudevan**

Christian Medical College & Hospital

**Shoma Rao**

Christian Medical College & Hospital

**Santosh Varughese**

Christian Medical College & Hospital

**Priscilla Rupali**

Christian Medical College & Hospital

**Maki Gina**

Henry Ford Health System

**Marcus Zervos**

Henry Ford Health System

**John Peter**

Christian Medical College & Hospital

**Balaji Veeraraghavan** (✉ [vbalaji@cmcvellore.ac.in](mailto:vbalaji@cmcvellore.ac.in))

Christian Medical College & Hospital

---

## Research Article

**Keywords:** MRSA, ST772, Integrated resistance plasmid, PVL, SCC*mec* V

**Posted Date:** January 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-141678/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Significant changes in the epidemiology of methicillin-resistant *Staphylococcus aureus* (MRSA) were recognised with the emergence of community-associated methicillin-resistant *Staphylococcus aureus*. However, studies on the molecular epidemiology and the genomic investigation of MRSA are limited in India. The aim of the study was to understand the molecular epidemiology of MRSA causing bloodstream infection and also to investigate the origin and evolution of ST772 *S. aureus* isolated from India.

## Results

SCCmec V (42%) was the predominant gene followed by SCCmec III (27%), SCCmec IV (13%), SCCmec I (7%) and SCCmec II (3%). We demonstrate the presence of multiple SCCmec types in 18 isolates. MLST analysis revealed ten different clonal complexes and three singletons. ST772 (27%), ST22 (19%) and ST239 (16%) were the predominant MRSA genotypes in causing bloodstream infection. The *spa* types were highly diverse. Phylogenetic analysis revealed that nearly three-fourth of the Indian ST772-SCCmec V isolates belongs to dominant (ST772-A2) and emerging subgroups (ST772-A3). A prophage  $\Phi$  IND772 carrying PVL toxin and a staphylococcal enterotoxin A was noticed in all isolates, except two. In addition, three distinct genomic islands (vSa-alpha, beta and gamma) were universally seen in all ST772 *S. aureus* genomes. A pattern of increasing antimicrobial resistance was noticed in the dominant and emerging subgroups. An integrated resistance plasmid encoding resistance clusters for beta-lactam (*blaZ*), macrolides (*mphC*, *msrA*), and aminoglycoside resistance (*aphA-III*, *sat-4*, *aadE*) was identified in all isolates, except four basal strains. ST772-SCCmec V was emerged on the Indian subcontinent in 1964 and diverged into a dominant subgroup in 1991. Furthermore, the expansion is likely to be associated with the acquisition of mobile genetic elements such as integrated resistance plasmid and SCCmec V (5C2) as well as the fixation of double serine mutation (S84L, S80Y) in the quinolone resistance determining region.

## Conclusions

ST772-SCCmec V has the multi-drug resistance trait of hospital-associated (HA) MRSA and the epidemiological characteristics of CA-MRSA. ST772 *S. aureus* have consistent virulence and resistance determinants which may results in successful survival in both community and hospital settings.

## Background

Methicillin-resistant *Staphylococcus aureus* (MRSA) is of significant clinical concern in causing both community-associated (CA) and healthcare-associated (HA) infections. MRSA was first reported in 1961 [1]. Methicillin resistance is conferred by the acquisition of staphylococcal cassette chromosome *mec* (SCC *mec*) carrying methicillin resistance determinant *mecA* [2]. The HA-MRSA strains carry the SCCmec

types I, II or III are often multi-drug resistant. Typically, the CA-MRSA strains carry the SCC*mec* types IV or V and a Panton-Valentine leukocidin (PVL) toxin are frequently susceptible to non-beta-lactam antibiotics [3]. Most of the HA-MRSA clones belong to the clonal complex (CC), CC5, CC8, CC22, CC30 and CC45 [4]. In contrast, CA-MRSA clones are geographically more diverse. In the Asia-Pacific region, five distinct sequence types (STs); ST59, ST30, ST72, ST8, and ST772 are the predominant and often spread to hospitals [5]. In India, MRSA accounts for 37% of *S. aureus* infections [6,7]. However, very few studies have reported the molecular epidemiology of *S. aureus*.

In early 2000, in India, most of the MRSA isolates belongs to the lineage ST239-SCC*mec* III and were restricted only to hospitals [8-10]. In 2004, a novel CA-methicillin susceptible *S. aureus* (MSSA) clone, ST772 was reported from a hospital in Bangladesh [11] and a ST772-SCC*mec* V (Bengal bay clone) isolated from the community and hospitals in India have been described [12]. During the same year, a variant of epidemic MRSA (EMRSA-15, ST22) carries SCC*mec* IV and a PVL have also been reported from India [12]. Later, ST772 and ST22 were predominant from all infection sites have been reported in India [13-18].

Studies have emphasized the contribution of mobile genetic elements (MGEs) in the emergence of ST772-SCC*mec* V as a multi-drug resistant and highly virulent lineage [19,20]. The notable feature of ST772-SCC*mec* V lineage including the presence of SCC*mec* V, PVL carrying  $\Phi$ IND772 prophage, three distinct pathogenicity islands (vSa-alpha, beta, and gamma) and an integrated resistance plasmid (IRP) encoding a cluster of resistant genes such as *blaZ* (beta-lactam resistance), *mphC* and *msrA* (macrolide resistance) *aphA* III and a partial *aadE* (aminoglycoside resistance) as well as *sat4* (streptothricin resistance). A recent phylogenetic analysis predicted that Indian subcontinent was an early hub for ST772-SCC*mec* V, from which it was disseminated across the globe [21]. However, only a limited number of isolates from India have been included in this previous study. Consequently, understanding the expansion of this lineage and the phylogenetic relationship of ST772 *S. aureus* isolates within India is crucial.

Studies focusing on *S. aureus* genomics remain largely unexplored in India. Further, the use of a phylogenomic approach to focus on the origin and spread of ST772 *S. aureus* in India has been hindered by the lack of whole genome data and sparse geographic sampling. Therefore, the on-going public health significance of this Bengal bay clone in India remains unknown. Here, we investigated the molecular epidemiology of MRSA causing bloodstream infection in the hospital setting between 2013 and 2019. Further, the genomes of ST772 *S. aureus* isolates from India were compared to understand the introduction and expansion of this multi-drug resistant (MDR) lineages. Bayesian phylogenetic analysis was performed for ST772 *S. aureus* isolates from India, to demonstrate the evolution of this lineage. In addition, virulence and resistance profile were analysed and correlated with the previously defined ST772 *S. aureus* subgroups.

## Results

## Molecular epidemiology of MRSA isolates

The distribution of *SCCmec* types and *spa* types among various STs of MRSA isolates is shown in Table 1. *SCCmec* V (42%) was the predominant gene followed by *SCCmec* III (27%), *SCCmec* IV (13%), *SCCmec* I (7%) and *SCCmec* II (3%). Notably, eighteen isolates had multiple *SCCmec* types and the most common combination was *SCCmec* III with *SCCmec* V. MLST analysis revealed ten different CC and three singletons (ST616, ST1598, ST1947) (Table 1). Three major CCs were CC1 (ST1, ST772), CC22 (ST22, ST217, ST636, ST896, ST1037, ST2371, ST3976) and CC8 (ST8, ST239, ST368, ST630, ST1803, ST3324) which accounted for 82% (n=192) of the isolates (Table 1). Thirty two distinct STs were identified, ST772 was the most common followed by ST22, ST239 and ST368 (Additional file 1: Figure S1). Although, STs were observed to be diverse, most of them were recognised as a SLV or DLV of the dominant STs (Additional file 1: Table S1). Thirty distinct *spa* types were identified (Table 1), of them t657 (34%) was predominant followed by t425 (18%), t037 (12%), t852 (10%), and t030 (8%). The remaining 25 different *spa* types were found in fewer than five isolates. Furthermore, PVL gene was identified in 66% (n=153) of the isolates and majority of them belongs to ST772 (34%, n=52) and ST22 (27%, n=42).

## Genomic analysis of *S. aureus* ST772-*SCCmec* V

### Phylogenetic analysis

A maximum likelihood phylogenetic tree was constructed using 32 MRSA genome (belongs to ST772-*SCCmec* V) sequenced in this study and 273 *S. aureus* genomes belongs to ST772 from previous studies (Additional file 1: Table S2). This collection comprised of MRSA (n=270) and MSSA (n=34), which also includes ten basal strains (4 MSSA, 6 MRSA). A genomic comparison was made between the sequenced ST772-*SCCmec* V MRSA (n=32) from Vellore and a collection of 273 *S. aureus* belongs to ST772 from 14 countries (including India) (Additional file 1: Figure S2). The earliest ST772 *S. aureus* isolates (n=7) in the entire global collection were from India and Bangladesh (reported in 2004). The phylogenetic comparison of ST772 *S. aureus* genome from India against the genomes available from other countries revealed that Indian isolates were distributed across all the previously defined subgroups (Basal strains, ST772-A1, ST772-A2 and ST772-A3) (Additional file 1: Figure S2). As expected, all the ST772-*SCCmec* V (n=32) MRSA genome sequenced in this study (isolated from the year 2013 onwards) belongs to the successful ST772-A clade (Fig 1). Nearly, half of the Indian isolates are clustered in the dominant subgroup ST772-A2 (56%) followed by ST772-A1 (23%) and ST772-A3 (15%).

Most of the isolates, 71% (n=61) carried *SCCmec* V (5C2) variant, while the composite *SCCmec* V (5C2 and 5) was seen in only two isolates (Fig 1). A transposon Tn4001 carrying aminoglycoside resistance gene (*aacA-aphD*) in *SCCmec* V was noticed in 73% (n=63) of the genomes. Furthermore, IRP was found in all the isolates, except four basal strains.

## Virulence and antimicrobial resistance

There is no significant variation in the distribution of virulence genes among subgroups (Fig 2). A 59.4 kb  $\Phi$  IND772 prophage carrying PVL-operon (*lukS* and *lukF*) and a staphylococcal enterotoxin A (*sea*) was noticed in all isolates, except two. In addition, three different genomic islands were identified in all isolates (including basal strains); a 41 kb, vSa-alpha genomic island containing an exotoxin cluster (*set6*, *set7*, *set8*, *set10*, *set11*, *set12*, *set13*, *set14*, *set15*), a 25.2 kb vSa-beta genomic island containing an enterotoxin gene cluster (*sec*, *seg*, *sei*, *sem*, *sen*, *seo*), and a 24.6 kb, vSa-gamma genomic island contains an extracellular fibrinogen binding protein (*ecb*), a fibrinogen binding protein (*fbp*), an alpha hemolysin (*hla*) and an exfoliative toxin A (*eta*).

Screening of AMR genes revealed variable gene distribution and QRDR mutations within and across subgroups (Fig 2). A pattern of increasing antimicrobial resistance was noticed in the dominant and emerging subgroups. Among AMR genes, the penicillin resistance encoding *blaZ* gene was identified in all the isolates. Furthermore, methicillin resistance encoding *mecA* gene was distributed across subgroups with the exception of basal strains (3 isolates) and ST772-A1 (14 isolates). In addition, aminoglycoside resistant are predominantly seen in isolates of dominant and emerging subgroups. The macrolide resistance encoding genes *mph* (C) and *msr*(A) were seen across the subgroups (ST772-A1, A2, A3), except basal strains. Analysis of quinolone resistance mechanism revealed S84L in *gyrA* with S80Y in *griA* were the most frequent double serine mutations (Fig 2). In the QRDR region, other mutations such as S84L in *gyrA*, S80Y in *griA* and S84L in *gyrA* with S80F in *griA* were also noticed.

### Bayesian time scaled phylogenetic analysis

The root to tip analysis revealed a strong correlation (correlation coefficient 0.6862 and R squared 0.4709) between the time of isolation and distance from root suggesting the temporal clock like evolution in the lineage (Additional file 1: Figure S3). Moreover, several demographic models were tested on the Indian ST772 *S. aureus* genomes (n=86) and compared their marginal likelihood estimation. This revealed that skyline model with strict clock was the best supported model (Additional file 1: Table S3). The mean substitution rate was estimated to be  $1.22 \times 10^{-6}$  substitution per site per year. Based on this rate, the most recent common ancestor (MRCA) of ST772 *S. aureus* was estimated as 1964 (95% highest posterior density [HPD], 1956-1970). Further, we then dated the acquisition and loss of various mobile genetic elements associated with antimicrobial resistance as well as QRDR mutations. This analysis suggests that IRP and SCCmec V (5C2) were gained in 1991 (95% HPD, 1988-1993). Bayesian hierarchical clustering using core SNPs segregated *S. aureus* ST772 isolates that had IRP and a variant of SCCmec V (5C2) into separate subgroups (Fig 3). The double serine mutation (S84L in *gyrA* with S80F in *griA*) was predicted to be replaced by S84L with S80Y in ~1998. This infers that the ST772 *S. aureus* multi-resistant subgroups are likely expanded rapidly after 1991 and has since sustained.

## Discussion

In India, PVL positive ST772-SCCmec V and ST22-SCCmec IV are reported as the dominant lineages in causing MRSA infections in community and hospitals [15-17]. This was confirmed in the present study.

Remarkably, in this study, a high degree of variability was seen among PVL positive ST22 MRSA, which affects not only the SCC*mec* marker but also the core genome. The presence of unusual SCC*mec* types in ST22 MRSA indicates an evolutionary genetic change and so demands deeper genomic investigation. Studies have demonstrated the existence of multiple SCC*mec* types in MRSA [22-24]. In the present study, 6% of isolates had multiple SCC*mec* types. This suggests that multiple transfer events of SCC*mec* types may have resulted in the integration of multiple *ccr* gene complexes in *S. aureus* [25].

Although, ST772 *S. aureus* has spread internationally, its expansion in India has not been previously studied. In the present study, we provide evidence for the long term persistence of ST772 *S. aureus* clone in India, with the recent dominance of ST772-SCC*mec* V(5C). Previously, three subgroups ST772-A1 (early branching subgroup), ST772-A2 (a dominant subgroup), and ST772-A3 (emerging subgroup) were defined by Steinig *et al* [21]. We also observed that early branching subgroup does no longer exist and replaced with the dominant and emerging subgroups as previously reported. In India, the evolution pattern of ST772-SCC*mec* V, is very similar to the previously reported global study [21]. Multi-drug resistance appears to be the consistent feature of ST772-SCC*mec* V, which greatly limit the oral treatment options (anti-staphylococcal  $\beta$ -lactams, erythromycin and trimethoprim-sulfamethoxazole) available for the management of community- acquired MRSA infections.

The core genome of ST772 *S. aureus* remains stable suggesting the strength of these genomes in adapting to evolutionary pressures for persistence. The substitution rate of ST772 *S. aureus* ( $1.22 \times 10^{-6}$ ) estimated in this study, which is comparable to the previous estimate of  $1.18 \times 10^{-6}$  substitutions/site/year [21]. Moreover, the mutation rate estimated in the present study is consistent with the previous estimates of globally dominant *S. aureus* lineages ST22 ( $1.3 \times 10^{-6}$ ), ST239 ( $1.6 \times 10^{-6}$ ) and USA 300 clone ( $1.3 \times 10^{-6}$ ) [26-28]. In this study, we demonstrate that the TMRCA of ST772 *S. aureus* lineage was emerged on the Indian subcontinent in 1960s, which is in agreement with the recent findings [21]. Our data shows temporal introduction of IRP with the sequential acquisition of SCC*mec* V (5C2) and accumulation of double serine mutation in 1991 resulted in the expansion of multi-drug resistant clones.

Multiple intercontinental spread from India to various countries in Europe, Australia and Asia have been reported [21]. In addition, outbreaks in the neonatal intensive care units in Ireland and a hospital in Norway have also documented [29,30]. However, intense endemic dissemination of this clone is not evident following intercontinental spread. Therefore, ST772-SCC*mec* V (Bengal bay clone) has a similar pattern of dissemination as with that of other CA-MRSA such as USA300 clone and the ST80-SCC*mec* IV lineage.

## Conclusions

In this study, ST772, ST22 and ST239 were observed as the predominant MRSA STs in causing bloodstream infections. Phylogenetic reconstruction demonstrated that ST772 *S. aureus* is evolved by the uptake of mobile genetic elements and the spread is likely depend on it. Interestingly, ST772 *S. aureus* have consistent virulence and resistant determinants that probably results in adaptation and has important implications for outbreak in hospitals. ST772-SCC*mec* V (Bengal bay clone) was emerged on

the Indian subcontinent in 1964 and diverged into a dominant clade in 1991. Moreover, acquisition of IRP, SCC*mec* V (5C2) and the double serine mutations associated with fluoroquinolone resistance are likely attributes for the success of this lineage. Multi-drug resistance in ST772-SCC*mec* V, greatly limits the use of oral antibiotics such as anti-staphylococcal beta-lactams, erythromycin and trimethoprim-sulfamethoxazole in the management of community acquired infections. Further studies are required to investigate the genetic basis for colonisation and transmission properties of hospital-adapted ST772-SCC*mec* V lineage.

## Methods

### Bacterial isolates

Non-repetitive MRSA (n=233, one isolate per patient) isolated from the blood cultures between 2013 and 2019 at the Department of Clinical Microbiology, Christian Medical College and hospital, Vellore, India were included in this study. MRSA isolates were identified using standard bacteriological methods including gram staining, culture and tube coagulase test. The study was approved by the Institutional Review Board of Christian Medical College, Vellore.

### Molecular characterisation of MRSA

SCC*mec* types were determined using multiplex PCR as described by Zhang *et al* and screened for the presence of Panton Valentine Leukocidin (PVL) encoding gene as described by McClure *et al* [31,32]. The subtypes of SCC*mec* IV were identified as described previously [33]. MLST of *S. aureus* was performed as defined by Enright *et al* [34]. STs and CC were assigned according to the MLST database (<https://pubmlst.org/saureus/>). *spa* typing was performed as described by Harmsen *et al* [35]. Assignment of the *spa* types were carried out using the public *spa* type database Ridom SpaServer ([www.spaserver.ridom.de](http://www.spaserver.ridom.de)).

### Genome sequencing and assembly

We selected a subset of 32 MRSA isolates possessing SCC*mec* V and belong to ST772 with a similar antimicrobial resistant profile (resistant to cefoxitin, gentamicin, erythromycin, ciprofloxacin and trimethoprim-sulfamethoxazole) for whole genome sequencing and comparative analysis. The selected subset contained isolates representing each year from 2013 to 2018.

### Ion Torrent sequencing

Whole genome sequencing was performed using IonTorrent PGM platform (Life Technologies, Carlsbad, CA). The sequencing library was prepared using 200ng of genomic DNA which was sheared to a fragment length of 400bp. Library preparation and sequencing were performed according to the manufacturer's instruction using Ion Plus Fragment Library Kit (Life Technologies, California, United

States). De novo assembly of the raw reads generated from ion torrent were assembled using AssemblerSPAdes software v4.4.0.1 in Torrent suite server version 4.4.3.

### **Publicly available genome data:**

The original ST772-SCC*mec* V global data of Steinig *et al* [21], Prabhakara *et al* [36] and Balakuntla *et al* [37] were supplemented with the genome of an additional 32 isolates sequenced in this study. Sequence reads of *S. aureus* ST772 isolates (n=273) were retrieved from the NCBI and European Nucleotide Archive (ENA) database and assembled with SPAdes v.3.12 [38]. DAR4145 (accession number CP010526) deposited from Mumbai was used as the reference genome to assemble the previously sequenced isolates (n=273) and the isolates sequenced in this study (n=32). Genomes of the isolates from veterinary sources used by Steinig *et al* were excluded in this study [21].

### **Single nucleotide polymorphism (SNP) based phylogenetic analysis**

Phylogenetic analysis of ST772 *S. aureus* was performed against the subset of publically available global genomes (n=273). All ST772 *S. aureus* sequences were mapped to the DAR4145 (ST772-SCC*mec* V; accession number CP010526) reference genome using BWA mem algorithm [39]. SNPs were called against the reference using Snippy v.4.5.1 [40] and filtered using FreeBayes [41]. Further, the recombination free core genome SNP alignment file generated by Gubbins was used as an input to compute the mean evolutionary rate of the genomes and time of the most recent common ancestor (MRCA) [42]. The RaXML-NG 0.5.0 was used to generate a maximum-likelihood (ML) tree based on 2612 (Indian genomes) and 6344 (global genomes) variant sites in the core genome [43]. The general time reversible model (GTR +  $\Gamma$ ) of nucleotide substitution was implemented with ascertainment bias correction (Lewis) and 100 bootstrap replicates as suggested by Steinig *et al* [21]. The tree with the highest likelihood was midpoint rooted and visualised with iTOL [44]. Phylogenetic analysis was explored for both global (including genomes from India) and only Indian *S. aureus* ST772 genome. The study isolates and the clades were assigned to previously described lineages [21].

### **Bayesian time scaled phylogenetic analysis**

The temporal phylogenetic structure was determined using Bayesian Evolutionary Analysis by Sampling Trees (BEAST) v1.10.4 package after the initial assessment of the temporal signal [45]. Sources and date of isolation of 305 ST772 *S. aureus* isolates used for temporal analysis, of them only Indian genomes (n=86) were used for BEAST analysis. The relationship between root to tip distance and the time of isolation were analyzed. Root-to-tip regressions (R2) may indicate the evolutionary rate and probable outliers among the isolates. Evolutionary rates and tree topologies were determined with the generalized time reversible (GTR) substitution models with gamma distributed among-site rate variation with four rate categories (GTR+ $\Gamma$ 4+I). The strict (assumes that branches in the tree evolves at constant evolution rate) and uncorrelated relaxed molecular clocks (assumes that each branch in the tree have independent evolution rate) were implemented with three demographic models, Constant size, Exponential and Bayesian skyline coalescent tree priors. The MCMC chain was run for 500 million steps with sampling of



20000 generations. We performed an additional duplicate run of 500 million steps and combined the log using LogCombiner v1. 8.2 when the effective sample size (ESS) estimates did not reach 200 [46]. The convergence and mixing of each run was manually inspected Tracer.v.1.7 to ensure that all the runs converged to an ESS of >200 [47]. To determine the best-fitting coalescent model to describe changes in effective population size over time, log marginal likelihoods were calculated using path sampling and stepping stone sampling methods. Finally, Bayes factor was used to determine the best fit model [48], a burn in of 20% was discarded from the runs and the maximum clade credibility (MCC) tree was generated using Treeannotator v.1.8.2 [49]. BEAST output was analyzed using Tracer v1.7, with uncertainty in parameter estimates reflected as the 95% highest probability density (HPD). The annotated phylogenetic tree was visualized using FigTree v.1.4.6 and the Bayesian skyline plot was reconstructed using Tracer.v.1.7 [50,51]. BEAST analysis was conducted separately for global (including genomes from India) and only Indian ST772 *S. aureus* genomes.

### Identification of virulence and resistance determinants

Genome data were analysed for the presence of virulence genes virulence factor database (VFDB) [52] and the acquired virulence genes using VirulenceFinder 2.0 (<https://cge.cbs.dtu.dk/services/VirulenceFinder/>). Acquired antimicrobial resistant (AMR) genes and chromosomal mutations in the quinolone resistance determining region of *gyrA* and *griA/parC* were identified using ResFinder 4.1 (<https://cge.cbs.dtu.dk/services/ResFinder/>). Insertion of the phage genome was identified using PHASTER [53]. Pathogenicity Islands were identified using the Pathogenicity island database [54].

## List Of Abbreviations

MRSA: Methicillin resistant *Staphylococcus aureus*; CA- community acquired; HA-hospital acquired; SCC – staphylococcal cassette chromosome; MLST: multi-locus sequence typing; *spa* – staphylococcal protein A; PVL: Panton Valentine Leukocidin; MGEs: mobile genetic elements; MDR: multi-drug resistance; IRP - integrated resistance plasmid; HPD: highest posterior density; TMRCA: the most recent common ancestor; SNP: single nucleotide polymorphism; BEAST: Bayesian Evolutionary Analysis by Sampling Trees; GTR: generalized time reversible;

## Declarations

**Ethics approval and consent to participate:** Not applicable

**Consent for publication:** Not applicable

**Availability of data and materials:** The genome shotgun sequencing data are deposited at NCBI/GenBank under the bioprojects PRJNA498584, PRJNA348458, PRJNA326798, PRJNA318841, PRJNA348465, PRJNA520811 and PRJNA348412.

**Competing interests:** The authors declare that they have no competing interests

**Funding:** None

**Authors' contribution:**

YDB - Conceptualization, methodology, formal analysis, data curation, writing Original Draft, writing-review and editing, Visualization. KV- Bioinformatic analysis, data curation. SVR, SV and PR – Investigation and critical review. MG, MZ and JVP- Investigation, data curation and critical review. BV – Conceptualization, critical review and supervision.

**Acknowledgements:** None

## References

1. Jevons M. Celbenin-resistant staphylococci. *Med. J* 1961; 1: 124–125.
2. Robinson DA, Enright MC. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 2003;47(12):3926-34.
3. Boyle-Vavra S, Daum RS. Community-acquired methicillin-resistant *Staphylococcus aureus*: the role of Panton-Valentine leukocidin. *Lab Invest* 2007;87(1):3-9.
4. Donker GA, Deurenberg RH, Driessen C, Sebastian S, Nys S, Stobberingh EE. The population structure of *Staphylococcus aureus* among general practice patients from The Netherlands. *Clin Microbiol Infect* 2009;15(2):137-43.
5. Huh K, Chung DR. Changing epidemiology of community-associated methicillin-resistant *Staphylococcus aureus* in the Asia-Pacific region. *Expert Rev Anti Infect Ther* 2016;14(11):1007-1022.
6. Rajkumar S, Sistla S, Manoharan M, Sugumar M, Nagasundaram N, Parija SC, Ray P, Bakthavatchalam YD, Veeraraghavan B, Kapil A, Walia K, Ohri VC. Prevalence and genetic mechanisms of antimicrobial resistance in *Staphylococcus* species: A multicentre report of the indian council of medical research antimicrobial resistance surveillance network. *Indian J Med Microbiol* 2017;35(1):53-60.
7. Veeraraghavan B, Walia K. Antimicrobial susceptibility profile & resistance mechanisms of Global Antimicrobial Resistance Surveillance System (GLASS) priority pathogens from India. *Indian J Med Res* 2019;149(2):87-96.
8. Arakere G, Nadig S, Swedberg G, Macaden R, Amarnath SK, Raghunath D. Genotyping of methicillin-resistant *Staphylococcus aureus* strains from two hospitals in Bangalore, South India. *J Clin Microbiol*. 2005;43(7):3198-202.
9. Arakere G, Nadig S, Ito T, Ma XX, Hiramatsu K. A novel type-III staphylococcal cassette chromosome mec (SCCmec) variant among Indian isolates of methicillin-resistant *Staphylococcus aureus*. *FEMS Microbiol Lett* 2009;292:141–8.

10. Gadepalli R, Dhawan B, Kapil A, Sreenivas V, Jais M, Gaiind R, Chaudhry R, Samantaray JC, Udo EE. Clinical and molecular characteristics of nosocomial methicillin-resistant *Staphylococcus aureus* skin and soft tissue isolates from three Indian hospitals. *J Hosp Infect* 2009;73:253–63.
11. Afroz S, Kobayashi N, Nagashima S, Alam MM, Hossain AB, Rahman MA, Islam MR, Lutfor AB, Muazzam N, Khan MA, Paul SK, Shamsuzzaman AK, Mahmud MC, Musa AK, Hossain MA. Genetic characterization of *Staphylococcus aureus* isolates carrying Panton-Valentine leukocidin genes in Bangladesh. *Jpn J Infect Dis* 2008;61(5):393-6.
12. Goering RV, Shawar RM, Scangarella NE, O'Hara FP, Amrine-Madsen H, West JM, Dalessandro M, Becker JA, Walsh SL, Miller LA, van Horn SF, Thomas ES, Twynholm ME. Molecular epidemiology of methicillin-resistant and methicillin susceptible *Staphylococcus aureus* isolates from global clinical trials. *J Clin Microbiol* 2008;46:842–7
13. Kulkarni GB, Pal PK, Veena Kumari HB, Goyal M, Kovoov JM, Nadig S, Arakere G. Community-acquired methicillin-resistant *Staphylococcus aureus* pyomyositis with myelitis: a rare occurrence with diverse presentation. *Neurol India* 2009;57:653–6.
14. Bouchiat C, El-Zeenni N, Chakrakodi B, Nagaraj S, Arakere G, Etienne J. Epidemiology of *Staphylococcus aureus* in Bangalore, India: emergence of the ST217 clone and high rate of resistance to erythromycin and ciprofloxacin in the community. *New Microbes New Infect* 2015;7:15–20.
15. Dhawan B, Rao C, Udo EE, Gadepalli R, Vishnubhatla S, Kapil A. Dissemination of methicillin-resistant *Staphylococcus aureus* SCCmec type IV and SCCmec type V epidemic clones in a tertiary hospital: challenge to infection control. *Epidemiol Infect* 2015;143:343–53.
16. 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.
17. 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.
18. Bakthavatchalam YD, Babu P, Munusamy E, Dwarakanathan HT, Rupali P, Zervos M, John Victor P, Veeraraghavan B. Genomic insights on heterogeneous resistance to vancomycin and teicoplanin in Methicillin-resistant *Staphylococcus aureus*: A first report from South India. *PLoS One* 2019; 14(12):e0227009.
19. Monecke S, Baier V, Coombs GW, Slickers P, Ziegler A, Ehricht R. Genome sequencing and molecular characterisation of *Staphylococcus aureus* ST772-MRSA-V, "Bengal Bay Clone". *BMC Res Notes* 2013;6:548.
20. Steinig EJ, Andersson P, Harris SR, Sarovich DS, Manoharan A, Coupland P, Holden MT, Parkhill J, Bentley SD, Robinson DA, Tong SY. Single-molecule sequencing reveals the molecular basis of multidrug-resistance in ST772 methicillin-resistant *Staphylococcus aureus*. *BMC Genomics* 2015;16:388.

21. Steinig EJ, Duchene S, Robinson DA, Monecke S, Yokoyama M, Laabei M, Slickers P, Andersson P, Williamson D, Kearns A, Goering RV, Dickson E, Ehricht R, Ip M, O'Sullivan MVN, Coombs GW, Petersen A, Brennan G, Shore AC, Coleman DC, Pantosti A, de Lencastre H, Westh H, Kobayashi N, Heffernan H, Strommenger B, Layer F, Weber S, Aamot HV, Skakni L, Peacock SJ, Sarovich D, Harris S, Parkhill J, Massey RC, Holden MTG, Bentley SD, Tong SYC. Evolution and Global Transmission of a Multidrug-Resistant, Community-Associated Methicillin-Resistant *Staphylococcus aureus* Lineage from the Indian Subcontinent. *mBio* 2019;10(6). pii: e01105-19.
22. Bhutia KO, Singh T, Adhikari L, Biswas S. Molecular characterization of community- & hospital-acquired methicillin-resistant & methicillin-sensitive *Staphylococcus aureus* isolates in Sikkim. *Indian J Med Res* 2015;142:330–335
23. Murugesan S, Perumal N, Mahalingam SP, Dilliappan SK, Krishnan P. Analysis of antibiotic resistance genes and its associated SCCmec types among nasal carriage of methicillin resistant coagulase negative *Staphylococci* from community settings, Chennai, Southern India. *J Clin Diagnostic Res* 2015;9:DC01–DC05
24. Nagasundaram N, Sistla S. Existence of multiple SCCmec elements in clinical isolates of methicillin-resistant *Staphylococcus aureus*. *J Med Microbiol* 2019;68(5):720-727.
25. Jamrozny DM, Harris SR, Mohamed N, Peacock SJ, Tan CY, Parkhill J, Anderson AS, Holden MTG. Pan-genomic perspective on the evolution of the *Staphylococcus aureus* USA300 epidemic. *Microb Genom* 2016 May 31;2(5):e000058.
26. Holden MT, Hsu LY, Kurt K, Weinert LA, Mather AE, Harris SR, Strommenger B, Layer F, Witte W, de Lencastre H, Skov R, Westh H, Zemlicková H, Coombs G, Kearns AM, Hill RL, Edgeworth J, Gould I, Gant V, Cooke J, Edwards GF, McAdam PR, Templeton KE, McCann A, Zhou Z, Castillo-Ramírez S, Feil EJ, Hudson LO, Enright MC, Balloux F, Aanensen DM, Spratt BG, Fitzgerald JR, Parkhill J, Achtman M, Bentley SD, Nübel U. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res* 2013;23(4):653-64.
27. Baines SL, Holt KE, Schultz MB, Seemann T, Howden BO, Jensen SO, van Hal SJ, Coombs GW, Firth N, Powell DR, Stinear TP, Howden BP. Convergent adaptation in the dominant global hospital clone ST239 of methicillin-resistant *Staphylococcus aureus*. *mBio*. 2015 Mar 3;6(2):e00080.
28. Glaser P, Martins-Simões P, Villain A, Barbier M, Tristan A, Bouchier C, Ma L, Bes M, Laurent F, Guillemot D, Wirth T, Vandenesch F. Demography and Intercontinental Spread of the USA300 Community-Acquired Methicillin-Resistant *Staphylococcus aureus* Lineage. *mBio*. 2016;7(1):e02183-15.
29. 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(3):841-7.
30. Blomfeldt A, Larssen KW, Moghen A, Haugum K, Steen TW, Jørgensen SB, Aamot HV. Bengal Bay clone ST772-MRSA-V outbreak: conserved clone causes investigation challenges. *J Hosp Infect* 2017

Mar;95(3):253-258.

31. Zhang K, McClure JA, Elsayed S, Louie T, Conly JM. Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome mec types I to V in methicillin-resistant *Staphylococcus aureus*. *J Clin Microbiol* 2005 ;43(10):5026-33.
32. McClure JA, Conly JM, Lau V, Elsayed S, Louie T, Hutchins W, Zhang K. Novel multiplex PCR assay for detection of the staphylococcal virulence marker Panton-Valentine leukocidin genes and simultaneous discrimination of methicillin-susceptible from -resistant staphylococci. *J Clin Microbiol*. 2006;44(3):1141-4.
33. Ghaznavi-Rad E, Nor Shamsudin M, Sekawi Z, van Belkum A, Neela V. A simplified multiplex PCR assay for fast and easy discrimination of globally distributed staphylococcal cassette chromosome mec types in methicillin-resistant *Staphylococcus aureus*. *J Med Microbiol* 2010;59(Pt 10):1135-9.
34. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol* 2000;38(3):1008-15.
35. Harmsen D, Claus H, Witte W, Rothgänger J, Claus H, Turnwald D, Vogel U. Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for spa repeat determination and database management. *J Clin Microbiol* 2003;41(12):5442-8.
36. Prabhakara S, Khedkar S, Shambat SM, Srinivasan R, Basu A, Norrby-Teglund A, Seshasayee AS, Arakere G. Genome sequencing unveils a novel sea enterotoxin-carrying PVL phage in *Staphylococcus aureus* ST772 from India. *PLoS One*. 2013;8, e60013.
37. 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.
38. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J comput Biol*. 2012;19(5):455-77
39. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;25(14):1754-60.
40. Seemann T. Snippy: fast bacterial variant calling from NGS reads. <https://github.com/tseemann/snippy>. Accessed 12 March 2020
41. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv* 2012. arXiv:1207.3907.
42. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic acids research* 2015;43(3):e15-.
43. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30(9):1312-3.

44. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research*. 2016;44(W1):W242-5.
45. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* 2007;7(1):214.
46. Rambaut A, Drummond AJ. LogCombiner v1. 8.2.8.2. 2015, <http://beast.bio.ed.ac.uk>.
47. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic biology*. 2018 Sep;67(5):901.
48. Li WL, Drummond AJ. Model averaging and Bayes factor calculation of relaxed molecular clocks in Bayesian phylogenetics. *Molecular biology and evolution*. 2012 Feb 1;29(2):751-61.
49. Helfrich P, Rieb E, Abrami G, Lücking A, Mehler A. TreeAnnotator: versatile visual annotation of hierarchical text relations. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* 2018 May.
50. Rambaut A. FigTree v1.4.2, A Graphical Viewer of Phylogenetic Trees. 2014, Available from <http://tree.bio.ed.ac.uk/software/figtree/>
51. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic biology* 2018;67(5):901.
52. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 2005; 33(suppl\_1):D325–D328.
53. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016; 44(W1):W16–W21.
54. Bertelli C, Laird MR, Williams KP; Simon Fraser University Research Computing Group, Lau BY, Hoad G, Winsor GL, Brinkman FSL. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 2017; 45(W1): W30-35.

## Tables

**Table 1:** Distribution of sequence types, *SCCmec* types and *spa* types among MRSA isolated from bloodstream infection.

Clonal complex (CC)/Singleton (n)	Sequence Type (ST) (n)	SCC <i>mec</i> type (n)	<i>spa</i> type (n)
CC1 (67)	ST1 (4)	IVa (1), V (3)	t425(3), t127(1)
	ST772 (63)	II (1), IVa (1), V (54), III,V (6), IVa,V (1)	t3387(2), t425(1), t458(2), t657 (58)
CC22 (63)	ST22 (44)	I (4), II (3), III (4), IVa(1), IVc(6), IVd(7), V(16), III,V (3)	t005(1), t021(2), t030(2), t1152(1), t13078(1), t1839(1), t2649(1), t309(1), t310(1), t3107(1), t3435(1), t425(7), t458(4), t474(1), t657(6), t852(13)
	ST217 (5)	I (2), IVa(1), V(1), III,V(1)	t309 (1), t657(3), t12562 (1)
	ST636 (1)	IVa	t304
	ST896 (1)	I	t030
	ST1037 (4)	I (1), V (2), III,V (1)	t425(2), t657(1), t12562(1)
	ST2371 (7)	I (4), IVd(2), V(1)	t005(1), t458(1), t6827(3), t852(2)
	ST3976(1)	IVd	t030
CC8 (62)	ST8 (2)	II (1), V (1)	t064(2)
	ST239 (37)	I (2), III (35)	t030(12), t037(25)
	ST368(20)	III (19), V (1)	t425(20)
	ST630 (1)	III	t030
	ST1803(1)	V	t064
	ST3324 (1)	III	t425
CC5(11)	ST6(7)	I (1), II (1), IVd (1), V (4)	t304(5), t657(1), t425(1)
	ST801(1)	IVa	t304
	ST1290(2)	IVa(1), IVd(1)	t131(1), t2649(1)
	ST1307(1)	IVd	t304
CC672 (11)	ST672 (8)	IVd (1), V (5), I,IVd (1), IVd,V(1)	t657(3), t852(5)

	ST361 (2)	III (1), V (1)	t852(2)
	ST2775(1)	IVd	t3841
CC72 (7)	ST72 (7)	III (1), V (4), III,V (2)	t425(4), t2743(2), t3092(1)
CC30 (2)	ST1482(1)	III,V	t021
	ST382(1)	V	t021
CC88 (1)	ST88 (1)	III,V	t2518
CC121 (2)	ST1358 (1)	V	t425
	ST649 (1)	V	t425
CC580(1)	ST580 (1)	II	t4615
Singleton (6)	ST616 (3)	I (1), IVa(2)	t030(1), t852(1), t005(1)
	ST1598(2)	III (1), V (1)	t3387(1), t3841(1)
	ST1947 (1)	IVc	t657

## Figures



Tree scale: 0.001

#### Subgroups

- Basal strains
- ST772-A1
- ST772-A2
- ST772-A3

#### MSSA/MRSA

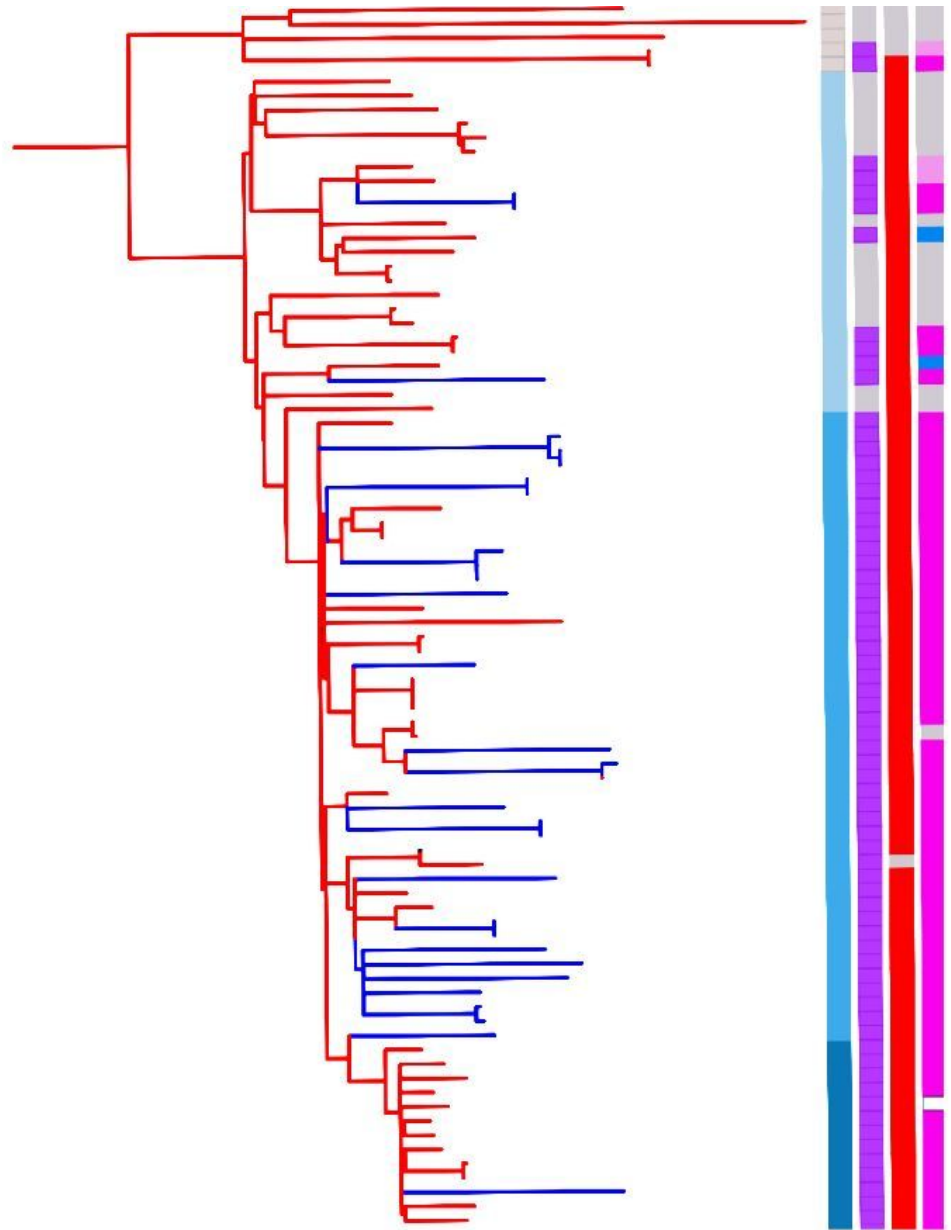
- MSSA
- MRSA

#### IRP

- Presence
- Absence

#### SCCmec types

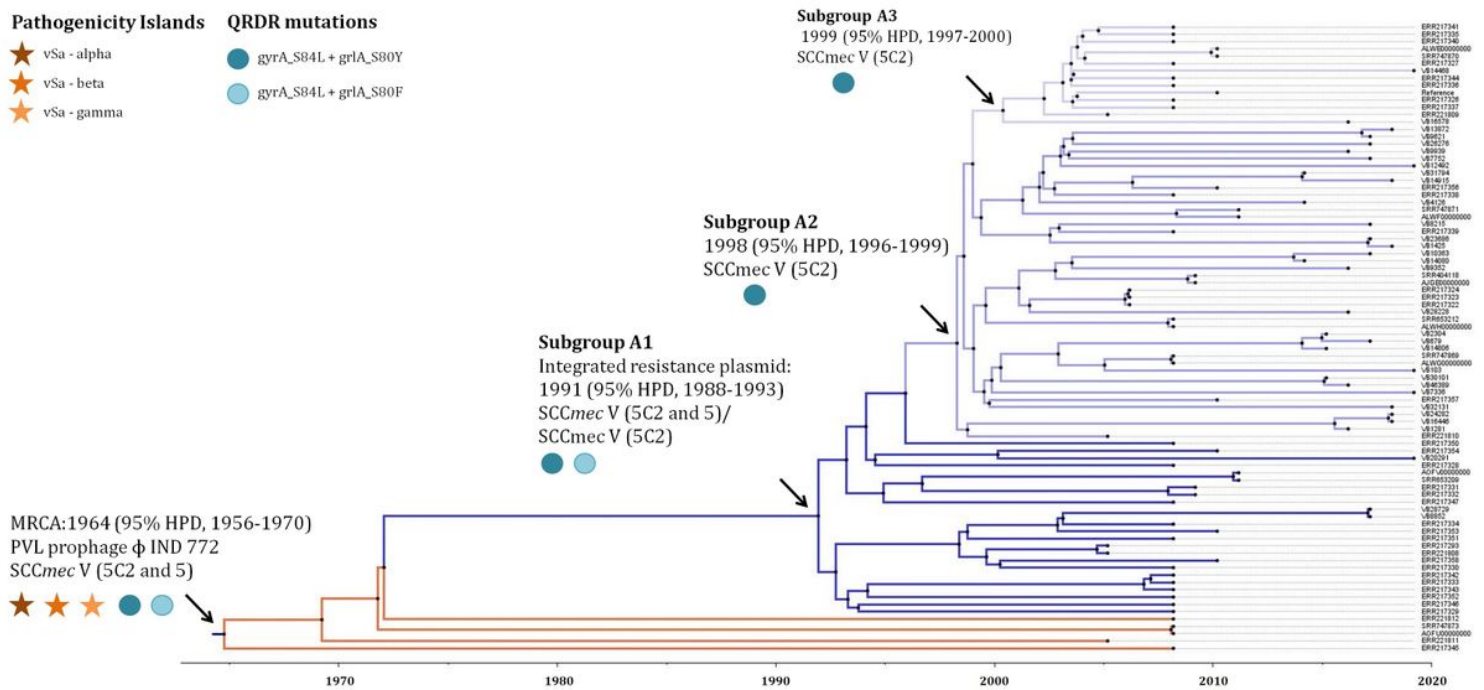
- SCCmec V (Bengal Bay)
- SCC negative
- PseudoSCCmec
- SCCmec VT (GR1)
- mecA



**Figure 1**

a) Maximum likelihood phylogenetic tree of 86 ST772 *S. aureus* isolated from India were mapped against DAR4145 reference genome. The first colour strip specifies the subgroups (ST772-A1, A2, A3). Isolates in the clades are colour coded (study isolates - blue in colour; other Indian isolates - red in colour). The bar indicate number of substitution per site.





**Figure 3**

Evolutionary analysis of 86 ST772 *S. aureus* isolated from India (2004-2018). The black arrow on the branches indicates the occurrence of apparent genetic events (including acquisition and loss of mobile genetic elements). The subgroups are in different colours.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigS3.jpeg](#)
- [Additionalfile.docx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS1.docx](#)
- [FigS1.jpeg](#)
- [FigS2.jpg](#)