

RESEARCH ARTICLE

# Genotypic Diversity of *Staphylococcus aureus* $\alpha$ -Hemolysin Gene (*hla*) and Its Association with Clonal Background: Implications for Vaccine Development

Meng Xiao<sup>1</sup>✉, Rui Zhao<sup>2</sup>✉, Qi Zhang<sup>1</sup>✉, Xin Fan<sup>1</sup>, Matthew V. N. O'Sullivan<sup>3,4</sup>, Dong-Fang Li<sup>5,6</sup>, Xin-Ying Wang<sup>5,6</sup>, Hong-Long Wu<sup>5,6</sup>, Fanrong Kong<sup>3</sup>, Ying-Chun Xu<sup>1\*</sup>

**1** Department of Clinical Laboratory, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China, **2** Department of Laboratory, Beijing Electric Power Hospital, Beijing, China, **3** Centre for Infectious Diseases and Microbiology Laboratory Services, ICPMR – Pathology West, Westmead Hospital, University of Sydney, Westmead, New South Wales, Australia, **4** Marie Bashir Institute for Infectious Diseases and Biosecurity, University of Sydney, Sydney, New South Wales, Australia, **5** Binhai Genomics Institute, BGI-Tianjin, BGI-shenzhen, Tianjin, China, **6** Tianjin Translational Genomics Center, BGI-Tianjin, BGI-shenzhen, Tianjin, China

✉ These authors contributed equally to this work.

\* [xycpumch@139.com](mailto:xycpumch@139.com)



OPEN ACCESS

**Citation:** Xiao M, Zhao R, Zhang Q, Fan X, O'Sullivan MVN, Li D-F, et al. (2016) Genotypic Diversity of *Staphylococcus aureus*  $\alpha$ -Hemolysin Gene (*hla*) and Its Association with Clonal Background: Implications for Vaccine Development. PLoS ONE 11(2): e0149112. doi:10.1371/journal.pone.0149112

**Editor:** D. Ashley Robinson, University of Mississippi Medical Center, UNITED STATES

**Received:** August 10, 2015

**Accepted:** January 27, 2016

**Published:** February 11, 2016

**Copyright:** © 2016 Xiao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All nucleotide sequence files are available from the NCBI GenBank database (accession numbers KT279554 to KT279561).

**Funding:** YCX received a National Research Special Fund for Public Welfare Industry of Health of China (grant number 201402001, <http://www.moh.gov.cn/qijiyis/s3577/201401/e9f3635e7acb47778225ccb729ffec62.shtml>) from the Ministry of Health of China. The funders had no role

## Abstract

The  $\alpha$ -hemolysin, encoded by the *hla* gene, is a major virulence factor in *S. aureus* infections. Changes in key amino acid residues of  $\alpha$ -hemolysin can result in reduction, or even loss, of toxicity. The aim of this study was to investigate the diversity of the *hla* gene sequence and the relationship of *hla* variants to the clonal background of *S. aureus* isolates. A total of 47 clinical isolates from China were used in this study, supplemented with *in silico* analysis of 318 well-characterized whole genome sequences from globally distributed isolates. A total of 28 *hla* genotypes were found, including three unique to isolates from China, 20 found only in the global genomes and five found in both. The *hla* genotype generally correlated with the clonal background, particularly the multilocus sequence type, but was not related to geographic origin, host source or methicillin-resistance phenotype. In addition, the *hla* gene showed greater diversity than the seven loci utilized in the MLST scheme for *S. aureus*. Our investigation has provided genetic data which may be useful for future studies of toxicity, immunogenicity and vaccine development.

## Introduction

*Staphylococcus aureus* is one of the leading human pathogens worldwide. It causes a broad range of diseases from superficial infections to life-threatening invasive diseases. Antimicrobial therapy is sometimes ineffective, owing to the development of antimicrobial-resistant strains, such as methicillin-resistant *S. aureus* (MRSA) [1]. *S. aureus* expresses various virulence

in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

factors, including a broad range of exotoxins. Some of these toxins, such as most of the phenol-soluble modulins and  $\alpha$ -hemolysin (also known as  $\alpha$ -toxin), are encoded in the core-genome; while others, such as Panton-Valentine leucocidin, are encoded by acquired mobile genetic elements [2, 3].

The  $\alpha$ -hemolysin, which belongs to a class of small  $\beta$ -barrel pore-forming cytotoxins, is a major virulence factor in *S. aureus* infections [2, 4]. It is encoded by the 960-bp *hla* gene, which is initially produced as a 319-residue precursor, then processed to a 293-residue (approximately 33 kDa) mature toxin [4, 5]. Previous studies have proven that changes in key amino acid residues of  $\alpha$ -hemolysin, such as a histidine substitution at amino acid 35, can result in reduction or even loss of virulence [6, 7].

In this study, we determined the genotypes of *hla* in 47 *S. aureus* isolates collected in China, and compared this with the clonal background of the isolates. For comparison, *hla* genotype and clonal background were also determined for well-characterized and published whole genome sequences of 318 global strains by *in silico* analysis.

## Methods

### Ethics

The study was approved by the Human Research Ethics Committee of Peking Union Medical College Hospital (No. S-263). Written consent was obtained from patients, and the study was carried out in accordance with approved guidelines.

### *Staphylococcus aureus* isolates

Isolates were collected from patients with *S. aureus* infections in Beijing Electric Power Hospital during January 2011 to June 2011. A total of 47 *S. aureus* isolates were collected consecutively, each from different individuals. The most common type of infection was pneumonia (n = 31, 66.0%), followed by soft tissue infection (n = 6, 12.8%), urinary tract infection (n = 6, 12.8%), bloodstream infection (n = 2, 4.3%), joint infection (n = 1, 2.1%), and gallbladder infection (n = 1, 2.1%).

### Molecular identification of isolates and *mecA* gene detection

DNA extraction of *S. aureus* isolates was performed as previously described [8]. A multiplex PCR was used for simultaneous amplification of 16S rRNA, *femA* and *mecA* genes for identification and differentiation of methicillin-susceptible *S. aureus* (MSSA) and MRSA isolates [9].

### Genotyping of the 960-bp *hla* gene

Primer pair of *hlaF1* (5' - TTAGCCGAAAAACATCATTTC-3') and *hlaR2* (5' - TTATTCCC GACGAAATTCCAA-3') was designed for amplification of the complete 960-bp *hla* gene encoding the  $\alpha$ -hemolysin precursor, which was made up of a 78-bp nucleotide sequence encoding the 26 aa signal peptide, followed by the 882-bp nucleotide sequence encoding the 293 aa mature  $\alpha$ -hemolysin.

Each PCR mix contained 12.5  $\mu$ L of 2 $\times$  EasyTaq PCR SuperMix (TransGen Biotech, Beijing, China), 2  $\mu$ L of DNA template, 0.5  $\mu$ M of each forward and reverse primer, and molecular biology grade water (TransGen Biotech) added to make a total volume of 25  $\mu$ L. PCR was performed as follows: initial denaturation at 95°C for 15 min, followed by 30 cycles of 94°C for 2 min, 55°C for 2 min, 72°C for 2 min, with a final extension at 72°C for 10 min. The products were sequenced in both directions using the inner primer pair *hlaF2* (5' - GAAGTTATCGGCTAAAGTTATAA-3')

and *hlaR1* (5' – CATAATTAATACCCTTTTCTC–3' ) on the DNA analyzer ABI 3730XL system (Applied Biosystems, Foster City, CA).

The obtained 960-bp *hla* gene sequences were compared to a wild-type reference sequence from *S. aureus* strain WOOD 46 (GenBank accession no. X01645) [10], aligned using CLC sequence viewer (version 7, QIAGEN Aarhus, Denmark) to detect single nucleotide polymorphisms (SNPs), and designate genotypes (the *hla* DNA sequence was identical for all isolates belonging to a given genotype). Further, the  $\alpha$ -hemolysin peptide sequences of isolates were deduced and aligned to determine the presence of amino acid substitutions.

## Assignment of clonal background

All isolates were analyzed by multilocus sequence typing (MLST) and *spa* typing using previously established methods [11, 12]. Assignment of related sequence types (STs) into clonal complexes (CCs) was conducted using eBURST [13]. In addition, all MRSA isolates were characterized by staphylococcal cassette chromosome *mec* (SCC*mec*) typing as described by Chen et al. [14]. *S. aureus* clones were named in the format of ST-*spa* type or ST-SCC*mec* type-*spa* type (e.g. ST5-t002 for a MSSA clone, or ST239-III-t030 for a MRSA clone).

## *In silico* analysis of published whole genome sequences

Because only limited number of *S. aureus* isolates was involved in the present study, to better define the *hla* gene diversity and association with clonal background amongst *S. aureus* with more different genetic background, 318 selected well-characterized published genomes derived from other geographic regions were further studied. The 318 selected published genomes included i) genetic information of all *S. aureus* isolates with complete assembled whole genome sequences as at July 8 2015 were obtained from the NCBI Genome database (70 isolates, S1 Table), and ii) *S. aureus* genome sequences from four previous publications (248 isolates, S2 Table) [15–18]. These genomes comprised examples from the major worldwide lineages of MRSA, e.g. CC8 (including USA-300), CC1 (including USA 400), CC22 (including EMRSA15), CC93 (including Queensland CA-MRSA), CC30 (including EMRSA16) and CC121 isolates.

The MLST STs and *hla* genotypes of the above isolates were determined by *in silico* mapping the paired-end reads of 318 isolates to seven gene loci sequences of *S. aureus* ST1 and wild-type *hla* reference sequence (GenBank accession no. X01645), respectively, using Burrows-Wheeler Alignment [19]. Base coverage of each position of genes was assessed using SAMtools mpileup packages (<http://samtools.sourceforge.net/mpileup.shtml>). The support number of reference base (ref) and alternative base (alt) are examined at each position in each strain. High quality SNPs were defined when SNPs satisfied the criteria of  $\text{alt}/(\text{alt}+\text{ref}) > 0.8$ . According to these SNPs, the ST and *hla* genotype for each isolate was identified.

MLST STs and *hla* genotypes were entered into BioNumerics software v7.5 (Applied Maths, Austin, TX) for minimum-spanning-tree analysis. The diversity of the sequences of the *hla* gene and seven genes utilized in the MLST scheme for *S. aureus* was analyzed by DnaSP (version 5.1, University of Barcelona, Spain).

## Results

### Molecular identification and detection of the *mecA* gene

All 47 isolates were confirmed as *S. aureus* by molecular methods. Thirty-three isolates (70.2%) were determined to be MRSA by detection of the *mecA* gene.

### Genotypes of the 960-bp *hla* gene

Amongst the 47 isolates from China, a total of eight *hla* genotypes (genotype 1 to 8) and six peptide sequence types were identified (Tables 1 and 2). Of the eight *hla* genotypes, genotype 1 was predominant (n = 27, 57% of 47 isolates), followed by genotype 3 (n = 7, 15%) and genotype 5 (n = 3, 6%). The remaining four genotypes were rare, with one or two isolates belonging to each (Table 1).

**Table 1. Nucleotide mutations and corresponding amino acid substitutions of the eight *S. aureus hla* genotypes identified from 47 isolates from China.**

Genotype	Nucleotide mutation position/corresponding amino acid substitution <sup>a</sup>																GenBank accession no.	
	-68	-22	-6&-5	144	177	399	438	453	479	606	624	669	708	765	777	824		864
Wild type <sup>a</sup>	G	G	G G	C	T	C	G	T	C	T	T	G	A	T	T	C	A	X01645.1
Genotype 1	A/	A/	AA/	T/	—	—	—	—	—	C/	—	—	—	C/	C/	T/	—	KT279554.1
	R (-23) H	Syn	G(-2) N	Syn						Syn				Syn	Syn	T275I		
Genotype 2	A/	A/	AA/	T/	—	—	—	—	—	C/	—	—	G/	C/	C/	T/	—	KT279555.1
	R (-23) H	Syn	G(-2) N	Syn						Syn			Syn	Syn	Syn	T275I		
Genotype 3	—	A/	AA/	—	—	—	—	—	—	—	G/	—	—	C/	—	—	—	KT279558.1
		Syn	G(-2) N								D208E			Syn				
Genotype 4	—	A/	AA/	—	A/	T/	C/	G/	—	C/	—	T/	—	C/	—	—	—	KT279556.1
		Syn	G(-2) N		Syn	Syn	Syn	Syn		Syn		Syn		Syn				
Genotype 5	—	A/	AA/	—	—	—	—	—	—	—	—	—	—	—	—	—	—	KT279559.1
		Syn	G(-2) N															
Genotype 6	—	A/	AA/	—	—	—	—	—	—	—	G/	—	—	C/	—	—	T/	KT279557.1
		Syn	G(-2) N								D208E			Syn			Syn	
Genotype 7	—	A/	AA/	T/	—	—	—	—	—	C/	—	—	—	C/	C/	T/	—	KT279561.1
		Syn	G(-2) N	Syn						Syn				Syn	Syn	T275I		
Genotype 8	—	A/	AA/	T/	—	—	—	—	Del <sup>b</sup>	C/	— <sup>b</sup>	— <sup>b</sup>	— <sup>b</sup>	C/	C/	T/	— <sup>b</sup>	KT279560.1
		Syn	G(-2) N	Syn						Syn <sup>b</sup>				Syn <sup>b</sup>	Syn <sup>b</sup>	T275I <sup>b</sup>		

Abbreviations: Syn, synonymous mutation; Del, deletion mutation; N/A, not applicable; “—”, no mutation.

<sup>a</sup>Nucleotide/peptide positions were designated relative to the first nucleotide/amino acid of the mature α-hemolysin (*S. aureus* strain WOOD 46, nucleotide sequence GenBank accession no. X01645).

<sup>b</sup>The deletion at nucleotide position 479 of genotype 8 resulted in peptide termination at residue position 164.

doi:10.1371/journal.pone.0149112.t001

**Table 2. Relationship between *hla* genotypes and molecular clone backgrounds of 47 *S. aureus* isolates from China.**

<i>hla</i> genotype	Molecular clone	Clonal Complex (CC)	No. of isolates
<b>Peptide sequence type 1</b>			
Genotype 1	ST239-MRSA-III-t030	CC8	13
Genotype 1	ST239-MRSA-III-t037	CC8	8
Genotype 1	ST239-MRSA-III-t2270	CC8	5
Genotype 1	ST239-MRSA-III-t459	CC8	1
Genotype 2	ST239-MRSA-III-t030	CC8	2
<b>Peptide sequence type 2</b>			
Genotype 3	ST5-MRSA-II-t002	CC5	3
Genotype 3	ST5-MSSA-t002	CC5	3
Genotype 3	ST25-MSSA-t078	CC25	1
<b>Peptide sequence type 3</b>			
Genotype 4	ST59-MRSA-IV-t437	CC59	1
Genotype 4	ST59-MSSA-t163	CC59	1
Genotype 5	ST1-MSSA-t127	CC1	3
<b>Peptide sequence type 4</b>			
Genotype 6	ST6-MSSA-t2467	CC6	1
Genotype 6	ST6-MSSA-t701	CC6	1
Genotype 6	ST7-MSSA-t091	CC7	1
Genotype 6	ST7-MSSA-t796	CC7	1
<b>Peptide sequence type 5</b>			
Genotype 7	ST188-MSSA-t189	CC1	1
<b>Peptide sequence type 6</b>			
Genotype 8	ST188-MSSA-t189	CC1	1

doi:10.1371/journal.pone.0149112.t002

Eighteen SNPs were found, including four on the signal peptide encoding portion (nucleotide positions -68, -22, -6 and -5 relative to the first nucleotide of mature  $\alpha$ -hemolysin portion), and 14 on the mature  $\alpha$ -hemolysin encoding sequence (nucleotide positions 144, 177, 399, 438, 453, 479, 606, 624, 669, 708, 765, 777, 824 and 864) (Table 1). The majority of the SNPs (12/18, 67%) were synonymous—only five (28%, nucleotide positions -68, -6, -5, 624 and 824) were nonsynonymous. Lastly, a deletion mutation was detected in nucleotide position 479 of genotype 8, presumably resulting in a prematurely terminating transcript (Table 1). Of the 18 SNPs identified, nucleotide position -68 was ST239-specific, positions 177, 399, 438, 453 and 669 were ST59-specific, position 765 was ST1-specific, whilst the other 11 were not lineage specific.

### Clonal background of *S. aureus* isolates

A total of six ST-SCC*mec-spa* types and nine ST-*spa* types were identified amongst 32 MRSA and 15 MSSA isolates from China, respectively (Table 2). 88% (29/33) of the MRSA isolates belonged to CC8, 9% (3/33) belonged to CC5, and 3% (1/33) belonged to CC59. The ST239-III-t030 clone comprised over half of the CC8 MRSA isolates (15/29, 52%), followed by ST239-III-t037 (8/29, 28%), ST239-III-t2270 (5/29, 17%) and ST239-MRSA-III-t459 (1/29, 4%). All three CC5 MRSA isolates were ST5-II-t002 (Table 2).

In comparison, the distribution of MSSA clones was more diverse, with no clone comprised of more than three isolates. ST5-t002 was represented by three MRSA and three MSSA isolates in this study, and ST59 was represented by one MRSA and one MSSA isolates with differing *spa* types. No other ST or CC was common to both MRSA and MSSA isolates (Table 2).

### Relationship between *hla* genotypes and clonal background

A strong correlation was observed between both *hla* genotypes and  $\alpha$ -hemolysin peptide sequence types, and the clonal background of isolates from China.

Of the eight *hla* genotypes, six were restricted to either MRSA (genotypes 1 and 2) or MSSA (genotypes 5, 6, 7 and 8) strains. The predominant MRSA clone, ST239-MRSA-III-t030, was represented by *hla* genotype 1 (13/15 isolates, 86.7%) or genotype 2 (2/15 isolates, 13.3%). All of the remaining ST239-MRSA-III isolates possessed *hla* genotype 1. Both genotype 1 and 2 *hla* genes encoded  $\alpha$ -hemolysin of peptide sequence type 1 (Table 2). *hla* genotypes 5 and 6 were found in one and four MSSA clones, respectively. Of the two ST188-MSSA-t189 isolates, one possessed *hla* genotype 7 and the other genotype 8, which differed by one deletion mutation at nucleotide position 479 (Table 1).

The remaining two *hla* genotypes, genotypes 2 and 3, were represented by both MRSA and MSSA isolates. Genotype 2 was found in six isolates of ST5-t002 (including three isolates each of ST5-MRSA-II-t002 and ST5-MSSA-t002) and one isolate of ST25-MSSA-t078. Genotype 3 was found in two ST59 isolates (one isolate each of ST59-MRSA-IV-t437 and ST59-MSSA-t163), and one isolate of ST1-MSSA-t127.

### The *in silico* analysis of 318 well-characterized genomes

Amongst 318 well-characterized *S. aureus* genomes, 25 *hla* genotypes were identified, including 20 genotypes not found in the 47 isolates from China (S1 and S2 Tables), for a total of 28 *hla* genotypes identified in this study. The SNPs identified for all clinical isolates and published genomes are summarized in S3 Table.

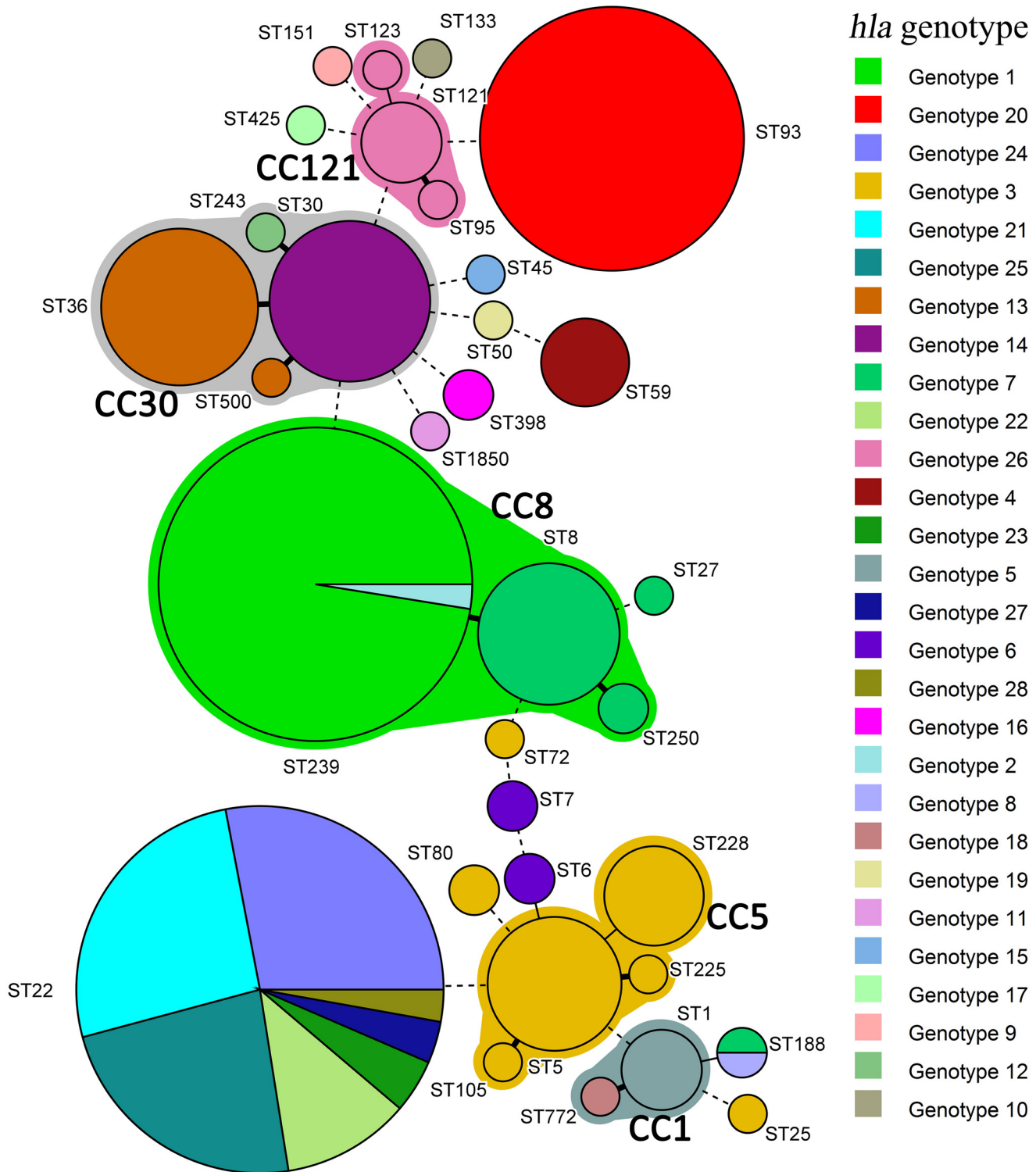
Substantial diversity amongst *S. aureus hla* gene was found. Compared to the seven loci utilized in the MLST, the *hla* gene had higher nucleotide diversity (0.0256 vs. 0.0042–0.0119), more haplotypes identified (28 vs. 10–19), greater haplotype diversity (0.899 vs. 0.517–0.804) and higher non-synonymous polymorphisms/ synonymous sites ratio (3.598 vs. 3.059–3.569) (Table 3). Of note, 79 of 107 ST22 *S. aureus* isolates and one ST188 isolate (*hla* genotype 8) had one to 16 deletion mutations in their *hla* sequences (S3 Table). In addition, all 20 ST36 isolates had a SNP C→T at sequence position 259, which resulted in a premature stop codon (S3 Table). These mutations would presumably inhibit production of the toxin protein.

In combining the 47 isolates from China and the 318 globally distributed genomes for composite analysis, it was again noted that *hla* genotype was closely related to the clonal background of the isolate, in particular the ST, with little association with the geographic origin, host source or methicillin-resistance phenotype (Table 2, S1 and S2 Tables and Fig 1). The minimum-spanning tree analysis of MLST data shown that for 30 of 33 STs identified in the present study, isolates belonging to the same ST shared a unique *hla* genotype, and ST22 was

**Table 3. Comparison of genetic diversity between the *hla* gene and seven gene loci utilized in the multilocus sequence typing scheme.**

Characters	Multilocus sequence typing loci							
	<i>hla</i>	<i>arcC</i>	<i>aroE</i>	<i>glpF</i>	<i>gmk</i>	<i>pta</i>	<i>tpi</i>	<i>yqiL</i>
Nucleotide diversity	0.0256	0.0083	0.0093	0.0042	0.0083	0.0079	0.0119	0.0073
No. of haplotypes	28	13	15	12	10	18	19	19
Haplotype diversity	0.899	0.760	0.796	0.517	0.726	0.804	0.800	0.799
No. of non-synonymous sites	730.09	350.45	356.18	350.45	327.30	362.89	312.13	400.19
No. of synonymous sites	202.91	105.55	99.82	114.55	91.70	111.11	89.87	115.81
No. of non-synonymous sites/ no. of synonymous sites	3.598	3.320	3.568	3.059	3.569	3.266	3.473	3.45574

doi:10.1371/journal.pone.0149112.t003



**Fig 1. Minimum-spanning tree of MLST data from 47 *S. aureus* clinical isolates collected in China and 318 whole genome sequences.** Circle sizes represent the number of isolates; circles are color coded by *hla* genotypes and labeled with the ST. Each circle represents a unique ST. Major clonal complexes (CCs), ie. CC8, CC5, CC30, CC121 and CC1, were labeled by green, brown, gray, pink and celadon colors, respectively. Thick solid lines, single-locus variant; Thin solid lines, double-locus variant; dashed lines,  $\geq 3$ -locus variant.

doi:10.1371/journal.pone.0149112.g001

the only sequence type that comprised more than three (seven in all) *hla* genotypes (Fig 1). In addition, STs belonging to the same CC frequently shared the same or close-related *hla* genotype, e.g. ST5, ST105, ST225 and ST228 of CC5 was comprised of 24 isolates belonging to *hla* genotype 3, ST8 and ST250 of CC8 was comprised of 18 isolates belonging to *hla* genotype 7, and ST95, ST121 and ST123 of CC121 was comprised of seven isolates belonging to *hla* genotype 26 (Fig 1).

Although the same *hla* genotype may be shared by unrelated STs, this was observed uncommonly. For instance, one isolate each that belonged to ST72 and ST25 isolates were *hla* genotype 3, which was mostly associated with CC5 *S. aureus* isolates. Likewise, one ST188 isolate was identified as *hla* genotype 25, which was mostly associated with non-ST239 CC8 isolates.

## Discussion

Infections due to antimicrobial-resistant pathogens are a growing problem all over the world. In developing countries like India and China, antimicrobial resistance is particularly prevalent, owing to previous unregulated overuse of antimicrobials [20, 21]. *S. aureus* is one of the commonest Gram-positive bacterial pathogens, and in many places, the majority of *S. aureus* infections are now caused by multidrug-resistant strains, including MRSA and vancomycin-resistant *S. aureus* (VRSA). Immunotherapies are now being investigated as an alternative therapeutic options for staphylococcal infections in the hope that these may avoid the selection pressure associated with the use of antimicrobials [3, 22].

The *S. aureus*  $\alpha$ -hemolysin was the first described of a family of bacterial pore-forming  $\beta$ -barrel toxins, which play an important role in the pathogenesis of staphylococcal disease [4, 23]. As such, it was chosen as a potential target for the development of vaccines to combat *S. aureus* infections, and positive results have been obtained in some preclinical trials targeting pneumonia and skin and soft tissue infections [23–26]. It has been noted that substitutions in amino acid residues may reduce the activity of  $\alpha$ -hemolysin. For instance, a  $\alpha$ -hemolysin mutant with a H35L substitution was found to have no hemolytic or lethal activity, despite retaining the ability to bind to target cells [6, 7]. The EMRSA-16 CC30 *S. aureus* isolates were another example. As observed in the present study, and as reported elsewhere, CC30-ST36 isolates had a SNP C→T at nucleotide sequence position 259, which resulted in a premature stop codon [27, 28]. It has been proven that CC30 isolates possessed this SNP had significantly reduced toxin production and decreased lethality in a mouse model [27]. These  $\alpha$ -hemolysin mutants could be considered as candidate immunogens in prototypic *S. aureus* vaccines [23–26, 29].

Despite this work, little has been described regarding the genetic polymorphism of the *hla* gene in *S. aureus*. This is an important consideration, since variation in  $\alpha$ -hemolysin peptide sequences could potentially lead to failure in antigen-antibody binding and thus compromise vaccine efficacy. In this study, we have illustrated the diversity of the *hla* gene in *S. aureus* and the relationship of *hla* sequence with clonal background, using 47 *S. aureus* clinical isolates from China supplemented with 318 well-characterized and globally distributed isolates with published whole genome sequences.

All ST239 from China were MRSA, and carried either genotype 1 or genotype 2 *hla*. These two genotypes differed by just one synonymous nucleotide mutation, (peptide sequence type 1). Amongst the 70 global *S. aureus* isolates with published genomes, seven isolates were ST239-MRSA-III, all of which also carried genotype 1 *hla*, regardless of the isolates' geographic origins. The ST239-MRSA-III clone has been reported largely to be hospital-acquired and widely disseminated in Brazil, Australia, New Zealand and many Asian countries in the past decade, although the prevalence different *spa* types within this clone (e.g. *spa* type t030 and



t037) vary in different regions [8, 30, 31]. In a previous genome-based phylogeographic analysis, it was determined that human movement played an important role in the global dissemination of ST239-MRSA-III [32]. Therefore, the consistent *hla* genotype of this clone across different regions is not surprising.

Interestingly, all of the 14 ST5 *S. aureus* isolates (six clinical isolates from China and eight global strains), including nine ST5-MRSA-II and five ST5-MSSA strains, possessed genotype 3 *hla*. Genotype 3 *hla* was also found in other CC5 *S. aureus* clones, including ST228-MRSA-I-t041 (n = 8), ST105-MRSA-II-t002 (n = 1) and ST225-MRSA-II-t003 (n = 1). Likewise all six ST59 isolates analyzed in this study, despite diverse methicillin-resistance phenotypes, SCCmec and *spa* types, carried genotype 4 *hla*. The ST59 lineage is primarily a community-acquired MRSA clone predominant in China and several other Asian countries [33]. These results again indicate that the *hla* genotype correlated closely with the ST.

Meanwhile, ST22 isolates shown significantly higher *hla* genotype diversity (comprised seven *hla* genotypes in all) than other *S. aureus* clones. Only occasional discrepancies between *hla* genotype and ST were observed. Future vaccine development will need to account for the influence of this diversity on vaccine effect.

## Conclusion

We have found substantial diversity amongst *S. aureus hla* gene and amino acid sequences. Strong correlations between *hla* genotypes and clonal background were found in *S. aureus*, regardless of the isolates' geographic origins and methicillin-resistance phenotype. Although the relative virulence of different *hla* genotypes remain undetermined, our investigation has provided some preliminary epidemiologic data which will be essential for future vaccine development.

## Supporting Information

**S1 Table. *In silico* analysis of *hla* genotype and clonal background of 70 complete assembled whole genome sequences of *S. aureus*.**

(DOCX)

**S2 Table. *In silico* analysis of *hla* genotype and clonal background of and 248 *S. aureus* genomes from four previous publications.**

(DOCX)

**S3 Table. Alignment of 365 *hla* sequences obtained from 47 China clinical isolates and 318 published genomes, and SNPs identified.**

(XLSX)

## Author Contributions

Conceived and designed the experiments: MX RZ YCX. Performed the experiments: MX QZ XF. Analyzed the data: MX MOS DFL XYW HLW. Contributed reagents/materials/analysis tools: MX FK. Wrote the paper: MX MOS FK.

## References

1. Stryjewski ME, Corey GR. Methicillin-resistant *Staphylococcus aureus*: an evolving pathogen. Clin Infect Dis. 2014; 58 Suppl 1:S10–9. doi: [10.1093/cid/cit613](https://doi.org/10.1093/cid/cit613) PMID: [24343827](https://pubmed.ncbi.nlm.nih.gov/24343827/).
2. Chen Y, Yeh AJ, Cheung GY, Villaruz AE, Tan VY, Joo HS, et al. Basis of virulence in a Pantone-Valentine leukocidin-negative community-associated methicillin-resistant *Staphylococcus aureus* strain. J Infect Dis. 2015; 211(3):472–80. doi: [10.1093/infdis/jiu462](https://doi.org/10.1093/infdis/jiu462) PMID: [25139021](https://pubmed.ncbi.nlm.nih.gov/25139021/).

3. Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat Rev Microbiol*. 2014; 12(1):49–62.
4. Berube BJ, Bubeck Wardenburg J. *Staphylococcus aureus* alpha-toxin: nearly a century of intrigue. *Toxins (Basel)*. 2013; 5(6):1140–66.
5. Walker B, Krishnasastri M, Zorn L, Kasianowicz J, Bayley H. Functional expression of the alpha-hemolysin of *Staphylococcus aureus* in intact *Escherichia coli* and in cell lysates. *J Biol Chem*. 1992; 267(15):10902–9. PMID: [1587866](#).
6. Menzies BE, Kernodle DS. Site-directed mutagenesis of the alpha-toxin gene of *Staphylococcus aureus*: role of histidines in toxin activity in vitro and in a murine model. *Infect Immun*. 1994; 62(5):1843–7. PMID: [8168947](#); PubMed Central PMCID: PMC186423.
7. Walker B, Bayley H. Key residues for membrane binding, oligomerization, and pore forming activity of staphylococcal alpha-hemolysin identified by cysteine scanning mutagenesis and targeted chemical modification. *J Biol Chem*. 1995; 270(39):23065–71. PMID: [7559447](#).
8. Xiao M, Wang H, Zhao Y, Mao LL, Brown M, Yu YS, et al. National surveillance of methicillin-resistant *Staphylococcus aureus* in China highlights a still-evolving epidemiology with 15 novel emerging multilocus sequence types. *J Clin Microbiol*. 2013; 51(11):3638–44. doi: [10.1128/JCM.01375-13](#) PMID: [23985906](#); PubMed Central PMCID: PMC3889745.
9. Al-Talib H, Yean CY, Al-Khateeb A, Hassan H, Singh KK, Al-Jashamy K, et al. A pentaplex PCR assay for the rapid detection of methicillin-resistant *Staphylococcus aureus* and Panton-Valentine Leucocidin. *BMC Microbiol*. 2009; 9:113. doi: [10.1186/1471-2180-9-113](#) PMID: [19476638](#); PubMed Central PMCID: PMC2700277.
10. Gray GS, Kehoe M. Primary sequence of the alpha-toxin gene from *Staphylococcus aureus* wood 46. *Infect Immun*. 1984; 46(2):615–8. PMID: [6500704](#); PubMed Central PMCID: PMC261580.
11. 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. PMID: [10698988](#); PubMed Central PMCID: PMC86325.
12. Koreen L, Ramaswamy SV, Graviss EA, Naidich S, Musser JM, Kreiswirth BN. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. *J Clin Microbiol*. 2004; 42(2):792–9. PMID: [14766855](#); PubMed Central PMCID: PMC344479.
13. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol*. 2004; 186(5):1518–30. PMID: [14973027](#); PubMed Central PMCID: PMC344416.
14. Chen L, Mediavilla JR, Oliveira DC, Willey BM, de Lencastre H, Kreiswirth BN. Multiplex real-time PCR for rapid *Staphylococcal* cassette chromosome *mec* typing. *J Clin Microbiol*. 2009; 47(11):3692–706. doi: [10.1128/JCM.00766-09](#) PMID: [19726600](#); PubMed Central PMCID: PMC2772623.
15. Stinear TP, Holt KE, Chua K, Stepnell J, Tuck KL, Coombs G, et al. Adaptive change inferred from genomic population analysis of the ST93 epidemic clone of community-associated methicillin-resistant *Staphylococcus aureus*. *Genome Biol Evol*. 2014; 6(2):366–78. doi: [10.1093/gbe/evu022](#) PMID: [24482534](#); PubMed Central PMCID: PMC3942038.
16. Holden MT, Hsu LY, 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(4):653–64. doi: [10.1101/gr.147710.112](#) PMID: [23299977](#); PubMed Central PMCID: PMC3613582.
17. McAdam PR, Templeton KE, Edwards GF, Holden MT, Feil EJ, Aanensen DM, et al. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A*. 2012; 109(23):9107–12. doi: [10.1073/pnas.1202869109](#) PMID: [22586109](#); PubMed Central PMCID: PMC3384211.
18. Kurt K, Rasigade JP, Laurent F, Goering RV, Zemlickova H, Machova I, et al. Subpopulations of *Staphylococcus aureus* clonal complex 121 are associated with distinct clinical entities. *PLoS One*. 2013; 8(3):e58155. doi: [10.1371/journal.pone.0058155](#) PMID: [23505464](#); PubMed Central PMCID: PMC3591430.
19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25(14):1754–60.
20. Mathur P, Singh S. Multidrug resistance in bacteria: a serious patient safety challenge for India. *J Lab Physicians*. 2013; 5(1):5–10. doi: [10.4103/0974-2727.115898](#) PMID: [24014960](#); PubMed Central PMCID: PMC3758706.
21. Xiao YH, Giske CG, Wei ZQ, Shen P, Heddini A, Li LJ. Epidemiology and characteristics of antimicrobial resistance in China. *Drug Resist Updat*. 2011; 14(4–5):236–50. doi: [10.1016/j.drug.2011.07.001](#) PMID: [21807550](#).

22. Ohlsen K, Lorenz U. Immunotherapeutic strategies to combat staphylococcal infections. *Int J Med Microbiol.* 2010; 300(6):402–10. doi: [10.1016/j.ijmm.2010.04.015](https://doi.org/10.1016/j.ijmm.2010.04.015) PMID: [20547101](https://pubmed.ncbi.nlm.nih.gov/20547101/).
23. Vandenesch F, Lina G, Henry T. *Staphylococcus aureus* hemolysins, bi-component leukocidins, and cytolytic peptides: a redundant arsenal of membrane-damaging virulence factors? *Front Cell Infect Microbiol.* 2012; 2:12. doi: [10.3389/fcimb.2012.00012](https://doi.org/10.3389/fcimb.2012.00012) PMID: [22919604](https://pubmed.ncbi.nlm.nih.gov/22919604/); PubMed Central PMCID: PMC3417661.
24. Kennedy AD, Bubeck Wardenburg J, Gardner DJ, Long D, Whitney AR, Braughton KR, et al. Targeting of alpha-hemolysin by active or passive immunization decreases severity of USA300 skin infection in a mouse model. *J Infect Dis.* 2010; 202(7):1050–8. doi: [10.1086/656043](https://doi.org/10.1086/656043) PMID: [20726702](https://pubmed.ncbi.nlm.nih.gov/20726702/); PubMed Central PMCID: PMC2945289.
25. Mocca CP, Brady RA, Burns DL. Role of antibodies in protection elicited by active vaccination with genetically inactivated alpha hemolysin in a mouse model of *Staphylococcus aureus* skin and soft tissue infections. *Clin Vaccine Immunol.* 2014; 21(5):622–7. doi: [10.1128/CVI.00051-14](https://doi.org/10.1128/CVI.00051-14) PMID: [24574539](https://pubmed.ncbi.nlm.nih.gov/24574539/); PubMed Central PMCID: PMC4018873.
26. Ragle BE, Bubeck Wardenburg J. Anti-alpha-hemolysin monoclonal antibodies mediate protection against *Staphylococcus aureus* pneumonia. *Infect Immun.* 2009; 77(7):2712–8. doi: [10.1128/IAI.00115-09](https://doi.org/10.1128/IAI.00115-09) PMID: [19380475](https://pubmed.ncbi.nlm.nih.gov/19380475/); PubMed Central PMCID: PMC2708543.
27. DeLeo FR, Kennedy AD, Chen L, Bubeck Wardenburg J, Kobayashi SD, Mathema B, et al. Molecular differentiation of historic phage-type 80/81 and contemporary epidemic *Staphylococcus aureus*. *Proc Natl Acad Sci U S A.* 2011; 108(44):18091–6. doi: [10.1073/pnas.1111084108](https://doi.org/10.1073/pnas.1111084108) PMID: [22025717](https://pubmed.ncbi.nlm.nih.gov/22025717/); PubMed Central PMCID: PMC3207694.
28. McGavin MJ, Arsic B, Nickerson NN. Evolutionary blueprint for host- and niche-adaptation in *Staphylococcus aureus* clonal complex CC30. *Front Cell Infect Microbiol.* 2012; 2:48. doi: [10.3389/fcimb.2012.00048](https://doi.org/10.3389/fcimb.2012.00048) PMID: [22919639](https://pubmed.ncbi.nlm.nih.gov/22919639/); PubMed Central PMCID: PMC3417553.
29. Menzies BE, Kernodle DS. Passive immunization with antiserum to a nontoxic alpha-toxin mutant from *Staphylococcus aureus* is protective in a murine model. *Infect Immun.* 1996; 64(5):1839–41. PMID: [8613399](https://pubmed.ncbi.nlm.nih.gov/8613399/); PubMed Central PMCID: PMC174000.
30. Carvalho KS, Mamizuka EM, Gontijo Filho PP. Methicillin/Oxacillin-resistant *Staphylococcus aureus* as a hospital and public health threat in Brazil. *Braz J Infect Dis.* 2010; 14(1):71–6. PMID: [20428658](https://pubmed.ncbi.nlm.nih.gov/20428658/).
31. Chen H, Liu Y, Jiang X, Chen M, Wang H. Rapid change of methicillin-resistant *Staphylococcus aureus* clones in a Chinese tertiary care hospital over a 15-year period. *Antimicrob Agents Chemother.* 2010; 54(5):1842–7. doi: [10.1128/AAC.01563-09](https://doi.org/10.1128/AAC.01563-09) PMID: [20176895](https://pubmed.ncbi.nlm.nih.gov/20176895/); PubMed Central PMCID: PMC2863666.
32. Gray RR, Tatem AJ, Johnson JA, Alekseyenko AV, Pybus OG, Suchard MA, et al. Testing spatiotemporal hypothesis of bacterial evolution using methicillin-resistant *Staphylococcus aureus* ST239 genome-wide data within a bayesian framework. *Mol Biol Evol.* 2011; 28(5):1593–603. doi: [10.1093/molbev/msq319](https://doi.org/10.1093/molbev/msq319) PMID: [21112962](https://pubmed.ncbi.nlm.nih.gov/21112962/); PubMed Central PMCID: PMC3115679.
33. Mediavilla JR, Chen L, Mathema B, Kreiswirth BN. Global epidemiology of community-associated methicillin resistant *Staphylococcus aureus* (CA-MRSA). *Curr Opin Microbiol.* 2012; 15(5):588–95.