Infection, Genetics and Evolution 59 (2018) 16–22

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

# Evolution of methicillin-resistant *Staphylococcus aureus*: Evidence of positive selection in a penicillin-binding protein (PBP) 2a coding gene *mecA*

Xiao-Yong Zhan[a,b,c,*], Qing-Yi Zhu[a,b]

[a] Guangzhou KingMed Center for Clinical Laboratory, Guangzhou 510300, China
[b] KingMed School of Laboratory Medicine, Guangzhou Medical University, Guangzhou 510300, China
[c] The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou 510080, China

ARTICLE INFO

*Keywords:*
*Staphylococcus aureus*
MRSA
*MecA*
PBP2a
Positive selection

ABSTRACT

Methicillin-resistant *Staphylococcus aureus* (*S. aureus*) (MRSA) represents more and more *S. aureus* infections. *MecA*, the novel coding gene of penicillin-binding protein (PBP) 2a of MRSA, is the key resistance factor of β-lactam, but little is known about the evolution of this gene. Given the crucial role of *mecA* in *S. aureus* physiology and β-lactam resistance, the selective forces may contribute to adaptation of the bacteria to the special environments such as its host or antibiotics. To understand the evolution of this gene, we screened GenBank database and analyzed *mecA* of 249 *S. aureus* strains. Twenty-nine unique alleles with 26 unique amino acid sequences were identified. Phylogenetic analysis showed three main groups of *mecA* in the global *S. aureus* strains. Analysis of these alleles using codon-substitution models (M8, M3, and M2a) and likelihood ratio tests (LRTs) of the codeML package and a random-effects likelihood (REL) method of HyPhy package for the site-specific ratio of nonsynonymous to synonymous substitution rates suggested that fourteen sites in the allosteric domain of PBP2a have been subjected to strong positive selection pressure. Mutations of two positive selection sites (N146K and E239K) were reported to be essential for ceftaroline- or L-695, 256-resistant. Further study indicated that the positive selection pressure might be more likely related to the host's inflammatory or immune response during *S. aureus* infection. Our studies provide the first evidence of positive Darwinian selection in the *mecA* of *S. aureus*, contributing to a better understanding of the adaptive mechanism of this bacterium.

## 1. Introduction

*Staphylococcus aureus* (*S. aureus*) is the main causative agent of hospital- and community-acquired infections worldwide (Peacock and Paterson, 2015). Methicillin-resistant *S. aureus* (MRSA) represents more and more *S. aureus* infections and acquired particular concern due to the increased prevalence of MRSA infections. The antibiotic resistance mechanism of MRSA is related to the acquisition of a set of genes that are induced and expressed on β-lactam exposure (Llarrull et al., 2009), in which penicillin-binding protein (PBP) 2a is the key resistance enzyme for virtually all β-lactam antibiotics (Stryjewski and Corey, 2014; Utsui and Yokota, 1985). PBP2a, a new PBP in concert with other PBPs plays an important role in cell wall biosynthesis of *S. aureus* including the catalyzation of peptide crosslinks between glycan chains of the cell wall (Giesbrecht et al., 1998; Pinho et al., 2001; Utsui and Yokota, 1985). β-Lactam antibiotics are substrate analogs of PBPs and inhibition of PBPs by β-lactams results in a weakened cell wall and cell lysis and death. The major difference of PBP2a with other PBPs is the reduced affinity for β-lactams, which makes the binding of β-lactam antibiotics weaken and remains active to allow cell wall synthesis at normally lethal β-lactam concentrations (Utsui and Yokota, 1985).

The PBP2a of MRSA consists of 668 amino acid residues. It is unlike any of the other PBPs produced by *S. aureus* (< 21% amino acid sequence identity) and the encoding gene, *mecA* (2007 bp) is relatively conserved among MRSA isolates (Garvey et al., 2016; Kuroda et al., 2001). Despite the prominent role of *mecA* in *S. aureus* pathogenicity and β-lactams resistance, little is known about the evolution of this gene. Rolo et al. studied the evolutionary steps leading to *mecA*-mediated β-lactam resistance in *staphylococci* (Rolo et al., 2017). They focused on the *mecA* in *Staphylococcus fleurettii*, the *mecA* precursor (*mecA1*) in *Staphylococcus sciuri*, the *mecA* and *mecA* homologue (*mecA2*) in *Staphylococcus vitulinus*, and *mecA* from *S. aureus* were not included in the study. Evolution of microorganisms during the infection results in part as a consequence of polymorphisms accumulated after selection pressure from the host's inflammatory or immune response (Fraser et al., 2005). Only a few examples of positive selection of bacterial genes were reported. Torres-Morquecho and colleagues reported that virulence genes of *cagA*, *babA*, and *oipA* of *Helicobacter pylori* (*H.*

* Corresponding author at: No.10, Luoxuan 3 Road, Guangzhou International Bio-Island, Guangzhou 510300, China.
*E-mail address:* tsinghan@126.com (X.-Y. Zhan).

https://doi.org/10.1016/j.meegid.2018.01.018
Received 24 August 2017; Received in revised form 19 January 2018; Accepted 22 January 2018

*pylori*) strains were placed under positive Darwinian selection (Torres-Morquecho et al., 2010). Stanhope et al. analyzed the PBPs of *Streptococcus pneumoniae* (*S. pneumoniae*) and found evidence of molecular adaptation with some positive selection sites correlating with amoxicillin susceptibility (Stanhope et al., 2008). However, these PBPs were from *S. pneumoniae* and only included PBP1a, 2b, and 2x. The primary purpose of this work is to study evidence of sites under positive selection in gene sequences of *mecA*, derived from a global *S. aureus* collection by employing molecular selection approaches. Our results showed that at least 14 amino acid sites of PBP2a were under positive selection and they were all located in the nonpenicillin-binding (nPB) domain which was associated with the allosteric control of PBP2a. Mutations of two positive selection sites (N146K and E239K) were reported to have special roles in the ceftaroline- or L-695, 256-resistant, and these sites may contribute to a stepwise increase in minimal inhibitory concentration (MICs) of antibiotics.

## 2. Materials and methods

### 2.1. S. aureus Strains and mecA sequences

*S. aureus mecA* was obtained by searching the GenBank nucleotide database (http://www.ncbi.nlm.nih.gov). TBLASTN program was used to collect sequences of *mecA* and *S. aureus* strain MRSA252 *mecA* was as queries (Holden et al., 2004). Finally, a total of 249 isolates were selected and manually verified for the present study. These strains were isolated during 1961 and 2015 in the different geographical locations including China, USA, UK, Canada, Germany, Ireland, Netherlands, Australia etc. Most of the strains were clinical isolates and a few were isolated from animals. The details of these strains including the GenBank accession numbers are shown in Table S1.

### 2.2. Analysis of mecA sequences

The *mecA* sequences were aligned by the MEGA7 package using Muscle (codons) parameters (Kumar et al., 2016). Allele type and DNA sequence polymorphism analyses were performed using DnaSP 5.10.01 (Librado and Rozas, 2009; Rozas, 2009). The PBP2a sequences of these strains were also aligned with the MEGA7 using Muscle parameters (Kumar et al., 2016), and polymorphism loci were analyzed.

An unrooted phylogenetic tree of the *mecA* was constructed using the Maximum Likelihood method of the MEGA 7 package (Kumar et al., 2016), based on the Kimura 2-parameter model (Kimura, 1980). Initial tree(s) was obtained by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. Bootstrap values were estimated by 1000 replications.

The aligned sequences of *mecA* were screened using RDP4 to detect intrinsic recombination (Martin et al., 2017), as to exclude the role of recombination in the *mecA* polymorphisms and demonstrate that CodeML and HyPhy package were fit for positive selection analysis. Six methods implemented in the program RDP4 were utilized to detect recombination. These methods were RDP (Martin and Rybicki, 2000), GENECONV (Padidam et al., 1999), BootScan (Martin et al., 2005), MaxChi (Smith, 1992), Chimaera (Posada, 2002) and SiScan (Gibbs et al., 2000). Potential recombination events were considered that identified by at least two methods according to Coscolla et al. report (Coscolla and Gonzalez-Candelas, 2009). Common settings for all methods were to consider sequences as linear, highest acceptable P-value was set at 0.05.

### 2.3. Analysis of positive selection

The selection pressure operating *mecA* was searched for using a Maximum Likelihood (ML) method. Analyses were performed using CodeML algorithm with site model implemented in PAML 4.7 (Yang, 1997, 2007). Three different nested models (M3 vs. M0, M2a vs. M1a, and M8 vs. M7) were compared and likelihood ratio tests (LRTs) were applied to select the one that best fitted the data. Positive selection was inferred when sites or codons with $\omega > 1$ were identified. When the LRT was significant ($P < 0.01$), Bayes empirical Bayes (BEB) methods (M8 model) and Naive Empirical Bayes (NEB) methods (M3 and M2a model) were used to identify amino acid residues that were likely to evolve under positive selection based on a posterior probability threshold of 0.95, and results from M8 model were taken as the standard one as Yang et al. recommended (Yang and Bielawski, 2000; Yang et al., 2005). M3 model was used for the frequency distribution of codon classes analysis as Yang et al. recommended, because it performed better than model M2a, and M8 model lacks this function. The HyPhy package was also utilized to identify the positive selection sites as a complement and random-effects likelihood (REL) method of the package was used (Kosakovsky Pond et al., 2011; Pond and Frost, 2005; Pond and Muse, 2005). A model selection procedure implemented in the HyPhy package was first run for finding the best nucleotide substitution bias model (Kosakovsky Pond and Frost, 2005). Positive selection was inferred when sites or codons with $\omega > 1$ and posterior probabilities were higher than 0.95.

### 2.4. Structural modeling

Three-dimensional structures were modeled using SWISS-MODEL (http://swissmodel.expasy.org), and PBP2a of strain 27r (PDB code: 1MWT) was used as the template (Guex et al., 2009; Lim and Strynadka, 2002). PBP2a from *S. aureus* contains an N-terminal transmembrane anchor, which can be removed without affecting the β-lactam binding, therefore soluble derivatives (residues 27–668) were used for structure determination (Lu et al., 1999). PBP2a of *S. aureus* strain BMS1, strain JCSC6943 and an artificial PBP2a sequence derived from the JCSC6943 with the positive selection sites substituted with possible amino acid residues (denoted JCSC6943*) were modeled. Protein secondary structure of the nPB domain of strain JCSC6943 and strain BMS1 was analyzed using ENDscript server (http://espript.ibcp.fr) (Robert and Gouet, 2014), on the basis of the structure obtained from SWISS-MODEL. Model quality was evaluated by QMEAN; the structure of the model was visualized by using PyMol (http://www. pymol.org).

## 3. Results and discussion

### 3.1. Characteristics of S. aureus strains, mecA sequences, and PBP2a

The *mecA* sequences of the 249 *S. aureus* strains in this study were clustered into 29 unique alleles (alleles were marked as representative strain name, Table S2). There is only one synonymous single-nucleotide polymorphism (SNP) site between alleles JCSC6943 and N315 (75A > C), alleles NCTC10442 and JCSC6945 (1830T > A), and alleles MRSA P126 and JCSC4610 (1461T > A). Other 23 alleles encode PBP2a with different amino acid sequences from each other; thus the 29 alleles corresponded to 26 different amino acid sequences (Table S3 and Fig. S1). We proposed that these sequences could represent almost all the *mecA* alleles up to now. Monecke et al. found only 17 alleles of *S. aureus mecA* by searching GenBank database in mid-2011 (Monecke et al., 2012). The increase of *mecA* alleles could be attributed to the complement of the database and new *S. aureus* strains discovery, as well as the *mecA* evolution. The number of polymorphic (segregating) sites of *mecA* was 54, while the single amino acid polymorphism loci of the PBP2a were 41, of which 29 loci were in the nPB domain (27 to 326 amino acid residues) of PBP2a, and 12 were in transpeptidase domain (327 to 668 amino acid residues, Fig. S1) (Lim and Strynadka, 2002; Otero et al., 2013).

The phylogenetic analysis revealed that *mecA* gene of *S. aureus* might consist of three main groups, and two alleles (BMS1 and X12) were evolutionary distinct from the others (Fig. 1A). Allele JCSC6943,