International Journal of Medical Microbiology 306 (2016) 686–696

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

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

Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of *Staphylococcus aureus*

Linyan Cao[a], Chun-Hui Gao[a], Jiade Zhu[a], Liping Zhao[a], Qingfa Wu[a], Min Li[b], Baolin Sun[a,*]

[a] *CAS Key Laboratory of Innate Immunity and Chronic Disease, School of Life Sciences and Medical Center, University of Science and Technology of China, Hefei, Anhui 230027, China*
[b] *Department of Laboratory Medicine, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai 200127, China*

ARTICLE INFO

Article history:
Received 26 May 2016
Received in revised form 25 July 2016
Accepted 29 August 2016

Keywords:
Staphylococcus aureus
Type III-A CRISPR-Cas system
SCC*mec*
Prophage

ABSTRACT

The CRISPR-Cas (clustered regularly interspaced short palindromic repeats [CRISPR]-CRISPR associated proteins [Cas]) system can provide prokaryote with immunity against invading mobile genetic elements (MGEs) such as phages and plasmids, which are the main sources of staphylococcal accessory genes. To date, only a few *Staphylococcus aureus* strains containing CRISPR-Cas systems have been identified, but no functional study in these strains has been reported. In this study, 6 clinical isolates of *S. aureus* with type III-A CRISPR-Cas systems were identified, and whole-genome sequencing and functional study were conducted subsequently. Genome sequence analysis revealed a close linkage between the CRISPR-Cas system and the staphylococcal cassette chromosome *mec* (SCC*mec*) element in five strains. Comparative sequence analysis showed that the type III-A repeats are conserved within staphylococci, despite of the decreased conservation in trailer-end repeats. Highly homologous sequences of some spacers were identified in staphylococcal MGEs, and partially complementary sequences of spacers were mostly found in the coding strand of lytic regions in staphylococcal phages. Transformation experiments showed that *S. aureus* type III-A CRISPR-Cas system can specifically prevent plasmid transfer in a transcription-dependent manner. Base paring between crRNA and target sequence, the endoribonuclease, and the Csm complex were proved to be necessary for type III-A CRISPR-Cas immunity.

1. Introduction

Staphylococcus aureus is a human pathogen that can cause a variety of diseases, from mild infections to life-threatening disorders (Lowy, 1998). Through horizontal gene transfer (HGT), *S. aureus* can acquire mobile genetic elements (MGEs) conferring high virulence and broad-spectrum antibiotic resistance, which enable it to respond to host immunity and survive antibacterial treatment (Lindsay, 2014). However, some of the MGEs, such as lytic phages, can bring detrimental threats. To deter such invaders, *S. aureus* has developed several antiviral defence systems, such as the restriction-modification system and the CRISPR-Cas (clustered regularly interspaced short palindromic repeats [CRISPR]-CRISPR associated proteins [Cas]) system (Westra et al., 2012).

The CRISPR sequence features direct repeats (DRs) and interspersed spacers. The repeats are conserved within a species and typically 24–48 bp in length, while the spacers vary from 26 to 72 bp, and most were derived from MGEs (Bolotin et al., 2005; Grissa et al., 2007; Mojica et al., 2005). According to the Cas proteins included, CRISPR-Cas systems are generally classified into two classes and five major types (type I to V) (Makarova et al., 2015). All these systems provide immunity against invading genetic elements via a three-stage process: adaptation, expression, and interference (van der Oost et al., 2014). The Cas proteins involved in interference can be guided to target sequence by specific crRNA via base paring. In type I and type II systems, PAM (protospacer adjacent motif) is important for target recognition in both adaptation and interference (Datsenko et al., 2012; Heler et al., 2015; Li et al., 2014; Sapranauskas et al., 2011; Sternberg et al., 2014).The PAM-proximal short sequence was identified as seed sequence, and base paring in this region is crucial for CRISPR-Cas immunity (Semenova et al., 2011; Wiedenheft et al., 2011). However, PAM is not found in type III system, and the complementarity between crRNA and

* Corresponding author.
E-mail address: sunb@ustc.edu.cn (B. Sun).

http://dx.doi.org/10.1016/j.ijmm.2016.08.005
1438-4221/

target in the upstream repeat region is responsible for self-nonself discrimination (Marraffini and Sontheimer, 2010).

Type III-A and III-B account for the majority of type III CRISPR-Cas systems (Makarova et al., 2015; Vestergaard et al., 2014), and their interference processes are mediated by the Csm and Crm complex, respectively (van der Oost et al., 2014). Unlike type III-B system, which mostly interferes with RNA (Hale et al., 2009; Zebec et al., 2014), type III-A system targets both DNA and RNA in a co-transcriptional way (Kazlauskiene et al., 2016; Samai et al., 2015). The nine Cas proteins in type III-A system are classified into three groups according to their functions (Makarova et al., 2015). Common among all CRISPR-Cas systems, the proteins Cas1 and Cas2 are involved in spacer acquisition (Wiedenheft et al., 2009). The meta-independent endoribonuclease Cas6 is essential for type III-A crRNA primary processing (Carte et al., 2008; Wakefield et al., 2015). The proteins Csm1 to Csm5 compose a Csm effector complex to directly interfere with target sequence (Hatoum-Aslan et al., 2013). The RNase Csm6 is not involved in DNA degradation, but provides an auxiliary RNA-targeting function to the Csm complex (Jiang et al., 2016; Niewoehner and Jinek, 2016).

To date, several *S. aureus* strains that harbour complete CRISPR-Cas systems have been identified (Golding et al., 2010; Holt et al., 2011; Kinnevey et al., 2013), but no functional study on these systems has been reported. *Staphylococcus epidermidis* has a type III-A CRISPR-Cas systems that also works in *S. aureus*, and it can specifically prevent plasmids conjugation and transformation, as well as phage infection (Goldberg et al., 2014; Marraffini and Sontheimer, 2008). However, the functions of *S. aureus* CRISPR-Cas system, especially in clinical isolates, need to be verified urgently. In this study, we identified 6 strains containing type III-A CRISPR-Cas systems through PCR-based screening among clinical isolates of *S. aureus*. Whole-genome sequencing and genome-wide sequence analysis of the strains were subsequently performed, followed by comparative sequence analysis of the CRISPR repeats and spacers. Transcription-dependent targeting of the type III-A CRISPR-Cas immunity in one strain was confirmed via plasmid challenge assay. The importance of base paring between crRNA and target sequence, and the contribution of each Cas proteins to immunity function were also investigated.

2. Results

2.1. Identification of six S. aureus strains harbouring type III-A CRISPR-Cas systems from 636 clinical isolates

To screen for *S. aureus* strains that harbour CRISPR-Cas system, a total of 636 clinical isolates were collected from four hospitals in Hefei and Shanghai, China (Supplementary Table S1 in the online version at DOI: 10.1016/j.ijmm.2016.08.005). The *S. aureus* strains were cultured overnight and then mixed by orthogonal design, and genomic DNAs were extracted from each of the mixture cultures. Degenerate primers for four *cas* genes (*cas1*, *cas2*, *cas6*, *cas10*) were designed according to three known staphylococcal type III-A systems (Supplementary Table S2 in the online version at DOI: 10.1016/j.ijmm.2016.08.005). Subsequently, extensive PCR experiments were conducted in parallel with these primers and the total DNA of clinical isolates as templates. Altogether, specific PCR products of all four *cas* genes were observed for 6 clinical isolates of *S. aureus*. (Fig. 1B). The CRISPR arrays were then amplified using additional primers (Supplementary Table S2 in the online version at DOI: 10.1016/j.ijmm.2016.08.005) and verified by sequencing. As a result, 6 strains harbouring type III-A CRISPR-Cas systems were identified and designated as *S. aureus* AH1, AH2, AH3, SH1, SH2, and SH3, respectively.

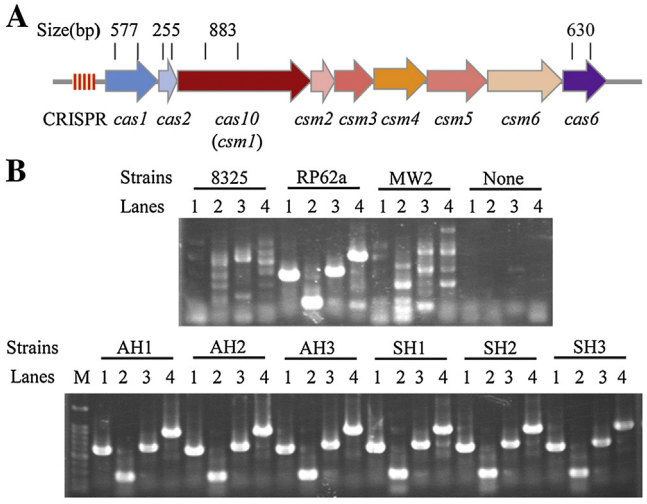

Fig. 1. Screening of *S. aureus* clinical isolates with type III-A *cas* genes. (A) The general organization of staphylococcal type III-A CRISPR-Cas system. The CRISPR array and *cas* genes were coloured and indicated on the bottom. The amplicon positions and sizes of *cas* genes were indicated on the top. (B) PCR results of 4 controls (upper panel) and 6 newly identified *cas*-containing clinical isolates (lower panel). M, marker; lane1, *cas1*; lane2, *cas2*; lane3, *cas6*; lane 4, *cas10*. The corresponding genomic DNA of indicated strains was used as template in each amplification. *S. aureus* strains NCTC8325 (8325) and MW2 were employed as negative controls, and *S. epidermidis* RP62a (RP62a) was employed as positive control. No template DNA was used in the blank control (None).

Table 1
Typing results of identified strains with type III-A CRISPR-Cas systems.

Strain	ST[a]	*Spa*	SCC*mec*
AH1	10-n-8-6-10-3-2	t1081	V
AH2	ST630	t4549	V
AH3	ST630	t4549	V
SH1	ST630	t4549	V
SH2	ST630	t4549	V
SH3	ST2250	t5078	None

[a] For a new ST type the ST profile was given, and n represents a new *aroe* allele.

2.2. Whole-genome sequencing of six clinical isolates of S. aureus with type III-A CRISPR-Cas systems

To dissect the genetic features of the identified strains, high-throughput genome sequencing was conducted to obtain their draft genome sequences. Overall, the genome sizes range from 2.78 Mbp to 2.92 Mbp, with a GC content of approximately 32.6%. As described in the Methods section, the coding sequences (CDS) were predicted and annotated. The Multilocus sequence typing (MLST), *spa* typing, and SCC*mec* typing were also conducted and shown in Table 1. The same typing results in four strains suggest that they are closely related, and consistently they also have similar CRISPR sequences. As in other *S. aureus* strains (Golding et al., 2010; Holt et al., 2011), the CRISPR-Cas loci in these strains were found in labile regions, mostly adjacent to their SCC*mec* elements, which is in agreement with the hypothesis that the CRISPR-Cas systems are mobile. However, strain SH3 has no SCC*mec*, implying that the transfer of the CRISPR-Cas system relies on other MGEs instead of SCC*mec*.

2.3. CRISPR repeats in staphylococcal type III-A CRISPR-Cas systems are conserved

To better understand the organization of type III-A CRISPR-Cas system in staphylococci, we investigated 15 sequenced staphylococcal strains with complete type III-A CRISPR-Cas systems

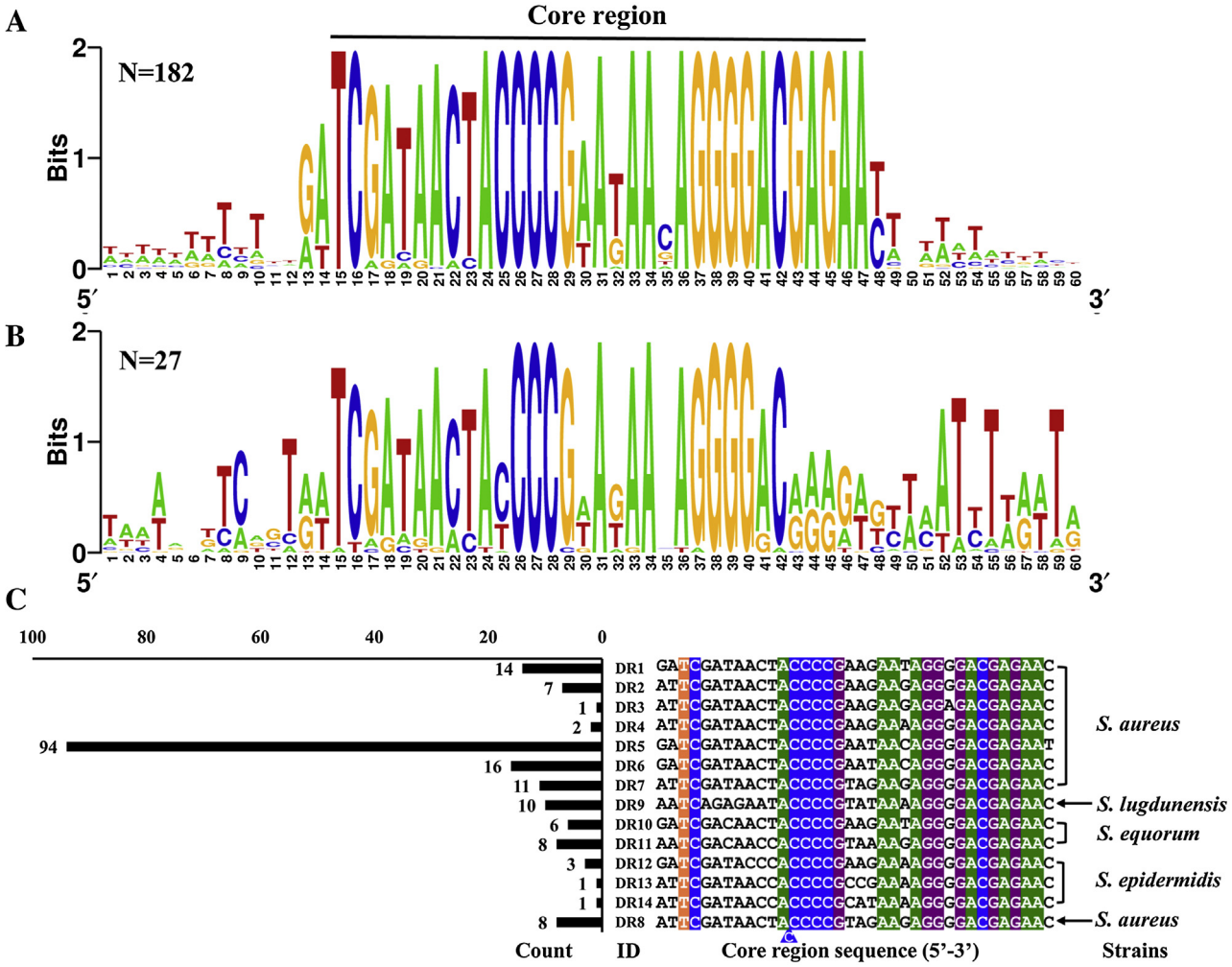

Fig. 2. Sequence conservation of staphylococcal type III-A CRISPR repeats. The leader-end and inner (A), or the trailer end repeats (B), including their upstream and downstream flanks, were aligned and shown as sequence logo. N indicates the number of repeats. (C) Classification of staphylococcal type III-A DRs according to the sequence. The count, ID, core region sequence, and strains were shown from left to right. An extra cytosine C of DR8 was shown alongside.

Table 2
Staphylococcal Type III-A CRISPR-Cas systems.

Strain[a]	Number of spacer[b]		Direct repeat
	Upstream	Downstream	
S. aureus AH1	3	N	DR1, DR2
S. aureus AH2	14	2	DR5, DR6, DR7, DR8
S. aureus AH3	15	2	DR5, DR6, DR7, DR8
S. aureus SH1	14	2	DR5, DR6, DR7, DR8
S. aureus SH2	15	2	DR5, DR6, DR7, DR8
S. aureus SH3	5	5	DR1, DR2, DR4
S. argenteus MSHR1132	6	4	DR1, DR2, DR3, DR4
S. aureus 08BA02176	15	3	DR5, DR6, DR7, DR8
S. aureus JS395	6	3	DR5, DR6, DR7, DR8
S.capitis CR01	15	2	DR5, DR6, DR7, DR8
S. schleiferi TSCC54	16	3	DR5, DR6, DR7, DR8
S. lugdunensis HKU09-01	N	6	DR9
S. lugdunensis N920143	N	4	DR9
S. equorum KS1039	6	8	DR10, DR11
S. epidermidis RP62A	3	2	DR12, DR13, DR14

[a] The generic name Staphylococcus was abbreviated to the initial capital letter S.
[b] N, no typical CRISPR array was characterized.

(Table 2). Interestingly, most strains were found to contain two CRISPR arrays with each one in upstream and downstream of the *cas* operons. Unexpectedly, *S. aureus* AH1 has a single repeat in the downstream of *cas* operon at the same positions of other downstream CRISPRs.

Analysis of the repeats indicates that the trailer-end repeats are not so conserved compared to the leader-end and inner repeats. Therefore, multiple sequence alignment of the 182 leader-end and inner repeats (including the single downstream repeat of *S. aureus* AH1) with their upstream and downstream flanking nucleotides was first performed, and the sequence conservation was visualized using WebLogo (Crooks et al., 2004). As shown in Fig. 2A, a highly conserved core region with a C cassette (four cytosines) and a G cassette (four guanines) was observed, and the downstream part of core region was more conserved compared with the upstream part, perhaps because the cleavage site for crRNA primary processing is near the G cassette (Hatoum-Aslan et al., 2011). A similar analysis was also performed on the 27 trailer-end repeats (Fig. 2B), in which the downstream conservation broke down. Obviously, the first cytosine of the C cassette in the trailer end repeats was frequently changed from C to T, which may affect the stability of the stem loop and the processing of the corresponding crRNAs.

The extremely random 12th and 50th sites seemed to be boundaries of the repeats (Fig. 2A). In this study, we defined the 36 bp sequence from 13th to 48th as the repeat, given that the 49th site was not conserved. According to the sequence, the 182 repeats were divided into fourteen subtypes (Fig. 2C). The seven

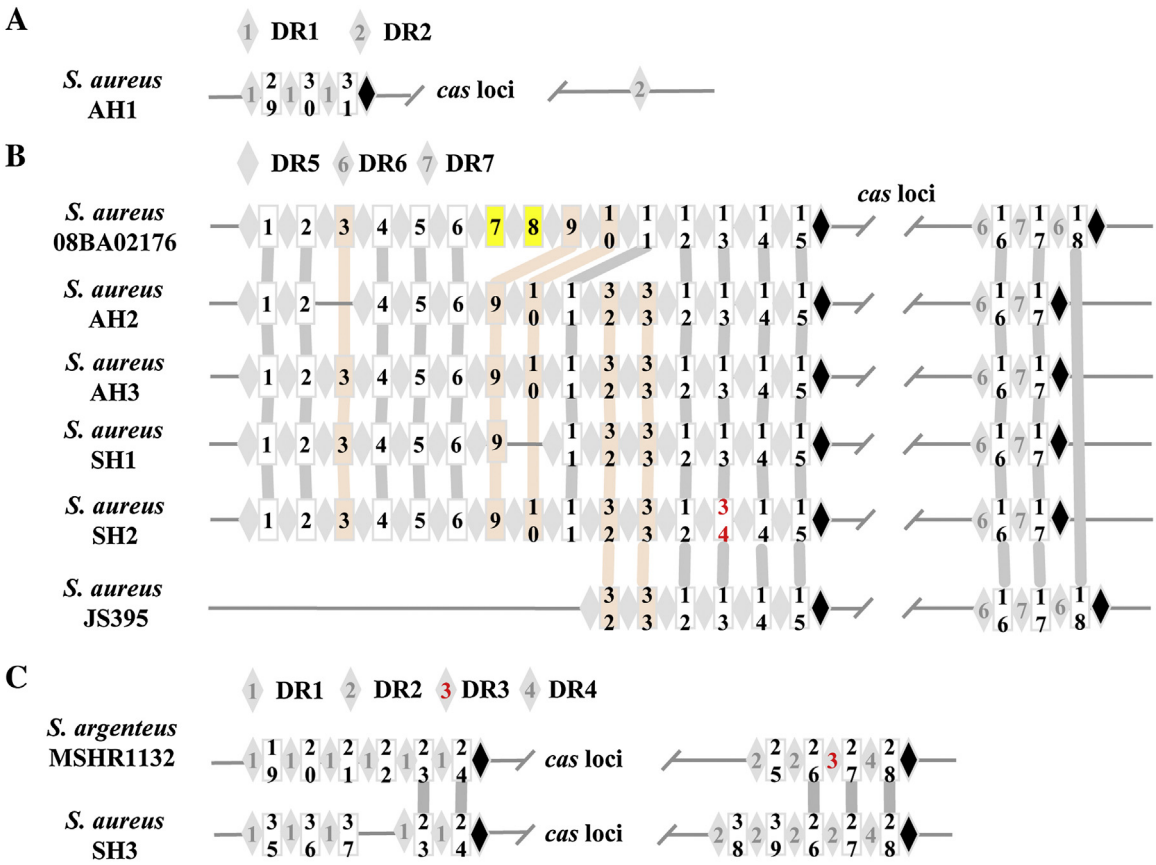

Fig. 3. The schematic diagrams of *S. aureus* type III-A CRISPR-Cas systems. (A) The schematic diagram of CRISPR-Cas system in *S. aureus* AH1. (B) The comparative schematic diagrams of CRISPR-Cas systems in *S. aureus* 08BA02176, AH2, AH3, SH1, SH2, and JS395. (C) The comparative schematic diagrams of CRISPR-Cas systems in *S. argenteus* MSHR1132 and *S. aureus* SH3. Repeats and spacers were shown as diamond and rectangle, respectively. The numeric identifier of repeat was given inside of diamond, and so did that of spacer but in vertical direction (from top to bottom). To enhance readability, colours were used, and the same spacers were linked and aligned vertically. S34 was linked with S13 as they only have one different nucleotide. The *cas* loci were folded since their differences were not taken into account in the analysis.

S. aureus DR subtypes have the same upstream variable sequence (GATAACT), implying that this region may be species specific. DR3 is located in the downstream CRISPR array of *Staphylococcus argenteus* MSHR1132 (Fig. 3C), and it appears to be a pseudo subtype of DR2 with only one different nucleotide. Because the mutation site is in the stem, we speculate that the immunity function of S27 could be at least partially impaired. In addition, it is interesting to note the existence of an extra cytosine (C) in DR8.

2.4. Comparative analysis of S. aureus CRISPR arrays

We further performed a comparative analysis of the *S. aureus* CRISPR arrays with respect to the spacers. The *S. argenteus* MSHR1132 was employed in the analysis because it shares similar CRISPR sequences with *S. aureus* SH3 (Fig. 3C). Besides, it was initially identified as an early-branching *S. aureus* lineage (Holt et al., 2011). In total, 39 unique spacers were characterized (Supplementary Table S3 in the online version at DOI: 10.1016/j.ijmm.2016.08.005), with some spacers that were common and had uniform relative positions among the newly identified and the known CRISPR arrays (Fig. 3B and C). Whereas, some spacers were only found in the newly identified CRISPR arrays, such as spacers S29–S31 in strain AH1 (Fig. 3A) and spacers S35–S39 in strain SH3 (Fig. 3C).

Additional findings were observed when we further analysed these arrays. Firstly, the CRISPR arrays are made up of repeat-spacer units, and the spacers pair up with specific repeats despite the diversity of repeats (Fig. 3). For example, all of the spacers S1 in these strains are paired with DR5, suggesting that the same spacers in these different CRISPRs may have the same origins. In addition, higher diversity of spacers has been observed in the upstream CRISPRs, where the leader-end units appear to be more changeable. As shown in Fig. 3B, some units are likely to be more dynamic than others. Compared with *S. aureus* 08BA02176, the four strains AH2, AH3, SH1, and SH2 all miss spacers S7, S8, and S18, and have S32 and S33 inserted between S11 and S12. And simultaneously, SH1 and AH2 miss S10 and S3, respectively. Meanwhile, strain JS395 misses many leader-end units that are located in upstream CRISPRs of other five strains. Consistently, *S. argenteus* MSHR1132 and *S. aureus* SH3 also share units in the trailer-end and differed in the leader-end (Fig. 3C). The larger size and diversity of the upstream CRISPR suggest that it may be more active in spacer acquisition, possibly due to the difference in its flanking sequences or the repeat sequences themselves.

2.5. Partially complementary target sequences of CRISPR spacers have strand bias

It is well recognized that the complementary sequence of spacer is the target of CRISPR-Cas immunity (Marraffini, 2015). Whereas, the CRISPR-Cas system may also perform other functions, in some cases through partial complementarity between spacer and its target sequence (Westra et al., 2014). Here we investigated the partially complementary sequences of the 39 spacers in the GenBank database, as described in the Methods section. As shown in Fig. 4A, high-quality partially complementary sequences of CRISPR spacers, which had normalized BLAST scores higher than 60, were

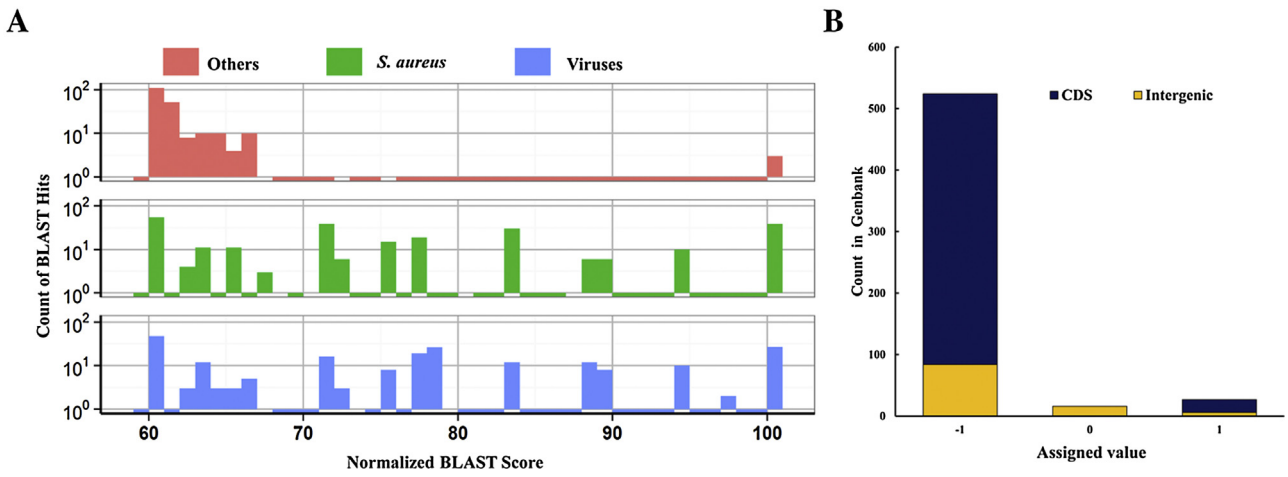

Fig. 4. Location of partially complementary sequences of CRISPR spacers. (A) Distribution of BLAST scores of 39 unique *S. aureus* CRISPR spacers. Blastn-short tasks were employed in this analysis with a cut-off of 1000 for the e-value. The spacers were used as queries to blast the NCBI nucleotide collection database (the 6 newly identified strains included). The results were divided into three parts: Virus/Phage hits under taxid 10239 (Virus), *S. aureus* hits under taxid 1280 (*S. aureus*) and the rest (Others). Histogram plots were generated for each spacer using R ggplot2 library (http://ggplot2.org/). (B) Location of partially complementary sequences of CRISPR spacers in GenBank. The assigned values were defined according to the transcription of the target sequences (−1, target sequence was in the template strand of transcribed region; 1, target sequence was in the coding strand of transcribed region; 0, target sequence was in intergenic region and its adjacent ORFs were in different orientations).

mainly distributed in *S. aureus* and viruses but not in other species. The protospacers have strand bias with most of them located in the template strand of target genes (Goldberg et al., 2014), i.e., the crRNAs could anneal to the transcript of target genes. To analyse the situation in the partially complementary sequences that were also regarded as CRISPR targets, we used an in-house Perl script to analyse their positions and transcription contexts. As shown in Fig. 4B, partially complementary sequences were found mainly in CDS regions. The assigned value according to their locations were dominantly -1, meaning that the sequences were mostly located in the template strand of transcribed regions. And the crRNAs could also anneal to corresponding target transcripts, although only with partial complementarity. Therefore, we perhaps can conclude that the partially complementary sequence of CRISPR spacer shares the same strand bias as the protospacer. In addition, the partially complementary sequences in staphylococcal prophages were mostly located in the lytic regions. For example, all of the partially complementary sequences of CRISPR spacers found in strain AH1 prophages were located in the lytic regions, away from the integrase genes (Supplementary Fig. S1 in the online version at DOI: 10.1016/j.ijmm.2016.08.005).

2.6. *Transcription-dependent DNA targeting of the* S. aureus *type III-A CRISPR-Cas immunity*

The type III-A CRISPR-Cas system in *S. aureus* AH1 has three spacers (Fig. 3A), and the first two spacers have matches in several staphylococcal phages. Spacer S29 matches to a region of gene coding for DNA polymerase, and S30 matches to the ORF region of a hypothetical gene (Supplementary Table S3 in the online version at DOI: 10.1016/j.ijmm.2016.08.005). To investigate whether the spacers of *S. aureus* AH1 have immunity function, we first employed an inducible plasmid pRMC2 (Corrigan and Foster, 2009). By inserting sequence of S29 in two orientations (direct and inverted) under the anhydrotetracycline-inducible promoter, the recombinant plasmids (pRS29D and pRS29I) were constructed. The inserted sequences in pRS29D and pRS29I were the same or complementary to S29, respectively. Then we transformed the plasmids into *S. aureus* AH1 wild type (WT) or CRISPR knockout cells by electroporation and checked the number of transformants. As a result, the AH1 WT strain gave nearly no transformant when transformed with pRS29SI, whereas it gave many transformants when using pRS29D, and the AH1 CRISPR knockout strain gave many transformants when using pRS29I (Fig. 5A). These results demonstrate that *S. aureus* type III-A CRISPR-Cas system performs DNA targeting in a transcription-dependent manner.

It's noteworthy that the above experiments were conducted under conditions with no inducer, meaning that the basal level transcription was enough for CRISPR-Cas immunity. To further confirm the transcription dependence of CRISPR-Cas immunity and determine the immunity function of the other two spacers, we employed a plasmid pALC-GFP in which GFP is constitutively expressed. By making sequences of S29 and S30 inserted in the 3′ end of *gfp* and sequence of S31 in the 5′ end, each in two orientations, we constructed the pASD plasmids (pAS29D, pAS30D, pAS31D) and the pASI plasmids (pAS29I, pAS30I, pAS31I). As to all the three spacers, no transformant was obtained from the AH1 WT strain when using the pASI plasmids, whereas many transformants were obtained when using the pASD plasmids, and many transformants were obtained from the AH1 CRISPR knockout strain when using pASI plasmids (Fig. 5B). These results confirmed that *S. aureus* AH1 type III-A CRISPR-Cas system has three functional spacers and provide antiplasmid immunity in a transcription-dependent manner.

While spacer S30 has perfect matching with its targets in sequenced phages, S29 has several mismatches with each of the potential target sequences in phages. To detect whether these mismatches would impair the CRISPR-Cas immunity effect, we constructed target plasmids with these potential target sequences integrated into pRMC2, and performed the transformation experiments. No transformant was obtained when the target plasmids were used to transform *S. aureus* AH1 WT strain (Supplementary Fig. S2 in the online version at DOI: 10.1016/j.ijmm.2016.08.005), suggesting that the mutations were tolerated by CRISPR-Cas immunity.

2.7. *The complementarity between crRNA and target in 5′ end is more important than that in 3′end*

Although the mismatches between S29 and the sequences in staphylococcal phages were tolerated by type III-A CRISPR-Cas immunity, we speculated that some mismatches would disrupt the immunity, and matching in some region may be more important. To investigate this, we stepwise mutated the perfectly matched target

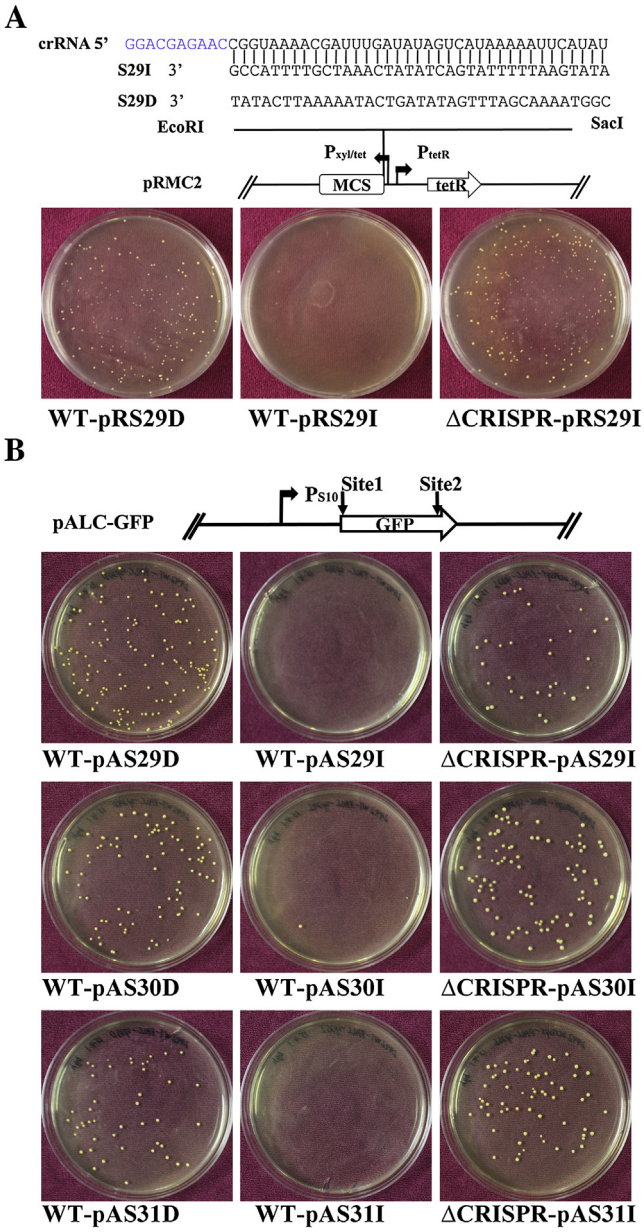

Fig. 5. The *S. aureus* CRISPR-Cas system provides immunity against plasmid in a transcription-dependent manner. (A) The transformation results of *S. aureus* AH1 WT or CRISPR knockout strains with target plasmids of S29. The base pairing between crRNA of S29 (with 5′ tag) and its target sequence were shown. The direct version (D) or inverted version (I) of spacer S29 were inserted into plasmid pRMC2 at indicated site. (B) The transformation results of *S. aureus* AH1 WT and CRISPR knockout strains with target plasmids of three spacers. Plasmid pALC-GFP was integrated with spacer sequences in directions D or I (S29, S30 in site2, S31 in site1). Plasmids used in these experiments were from strain RN4220. Three repeat experiments were performed to confirm the results, and the representative pictures were shown.

sequence from either ends by substituting the original nucleotides with their complementary nucleotides (Supplementary Table S4 in the online version at DOI: 10.1016/j.ijmm.2016.08.005). The plasmid containing perfectly matched target sequence was named pRproS29I, and the mutant plasmids were classified into three groups, "5PM", "3PM", and "5P_3PM". The "5PM" group introduced mismatches in the 5′ end region of S29, the "3PM" group introduced mismatches in the 3′ end, and the "5P_3PM" group introduced mismatches in both ends. Then transformation experiments were performed to detect their effects on immunity. When mutations in the 5′ end (Fig. 6A) were introduced, three or more consecutive nucleotide mutations in the first five sites (5P1-3M, 5P2-4M, 5P3-5M, 5P1-4M, 5P1-M) completely disrupted the immunity, but three consecutive nucleotide mutations in 6th-10th sites (5P4-6M, 5P5-7M, 5P6-8M, 5P7-9M, 5P8-10M) had no effect on immunity. In addition, two nucleotide mutations containing the second site (5P1-2M, and 5P2-3M) partially impaired the immunity, as less and smaller colonies were observed. When mutations in the 3′ end (Fig. 6B) were introduced, 13 consecutive nucleotide mutation was needed to completely disrupt the immunity (3P1–13M), 12 consecutive nucleotide mutation resulted in decreased immunity efficiency, and less consecutive nucleotide mutations had not effect. Besides, mutations introduced in both ends simultaneously exhibited synergy effect (Fig. 6C), because one or two nucleotide mutations in the 5′ end enhanced the effect of the mutations in the 3′ end (5P1_3P1-10M, 5P1_3P1-11M, 5P1-2_3P1-3M, 5P1-2_3P1-6M, 5P1-2_3P1-9M). Based on above results, we concluded that complementarity between crRNA and its target in the 5′ end is more important than that in the 3′ end, and mutations in both ends have synergy effect on immunity.

2.8. The Csm complex and Cas6 are required for antiplasmid CRISPR-Cas immunity

The type III-A CRISPR-Cas system of *S. aureus* is composed of one or more clusters of CRISPR arrays and nine Cas proteins. The participation of Cas proteins in CRISPR-Cas immunity against plasmid conjugation was investigated in *S. epidermidis* (Hatoum-Aslan et al., 2014), and sequence alignment revealed high homology of the Cas proteins between *S. aureus* AH1 and *S. epidermidis* RP62a (Supplementary Table S5 in the online version at DOI: 10.1016/j.ijmm.2016.08.005). To study the contribution of each Cas protein to the process of interference in *S. aureus*, we conducted in-frame deletions on the *cas*/*csm* genes in strain AH1, and checked their effects by transforming the corresponding mutant strains with plasmids pRS29D and pRS29I. As shown in Fig. 7, deletion of *csm1*, *csm2*, *csm3*, *csm4* or *cas6* completely disrupted the immunity, while deletion of *cas1* or *cas1-cas2* had no effect on the immunity function. In addition, deletion of *csm5* or *csm*6 partially impaired the immunity, as less and smaller colonies were obtained from their knockout strains when transformed with target plasmids.

3. Discussion

The CRISPR-Cas system is a double-edged sword to its host since it can both provide protection against invaders and hinder the gain of beneficial accessary genes (Bondy-Denomy and Davidson, 2014). It has been found in about 90% of archaea and about 50% of bacteria (Makarova et al., 2015). The relative less distribution in bacteria may be attributed to the existence of other defence systems and the important roles of accessary genes (Westra et al., 2012). Four ST630 strains in this study have similar CRISPR-Cas systems with that of ST398 isolate *S. aureus* 08BA02176. The livestock-associated (LA) ST398 was supposed to encode less virulence determinants and have low incidence in human in Canada (Golding et al., 2012). Whereas, the study on LA-MRSA CC398 in Germany revealed a high cytotoxic potential (Ballhausen et al., 2014), and ST398 is emerging as a community-associated lineage in China (Wang et al., 2016). The CRISPR-Cas systems seem to be enriched in some lineages, and the low occurrence detected in this study could partially be attributed to the fact that those lineages are not popular in China. Since the six strains in this study only represent three lineages, it is reasonable to speculate that *S. aureus* strains with divergent repeats should exist. That ST630 and ST398 harbour the similar CRISPR-Cas system is interesting, as ST630 is also an emerging lineage in China (Chao

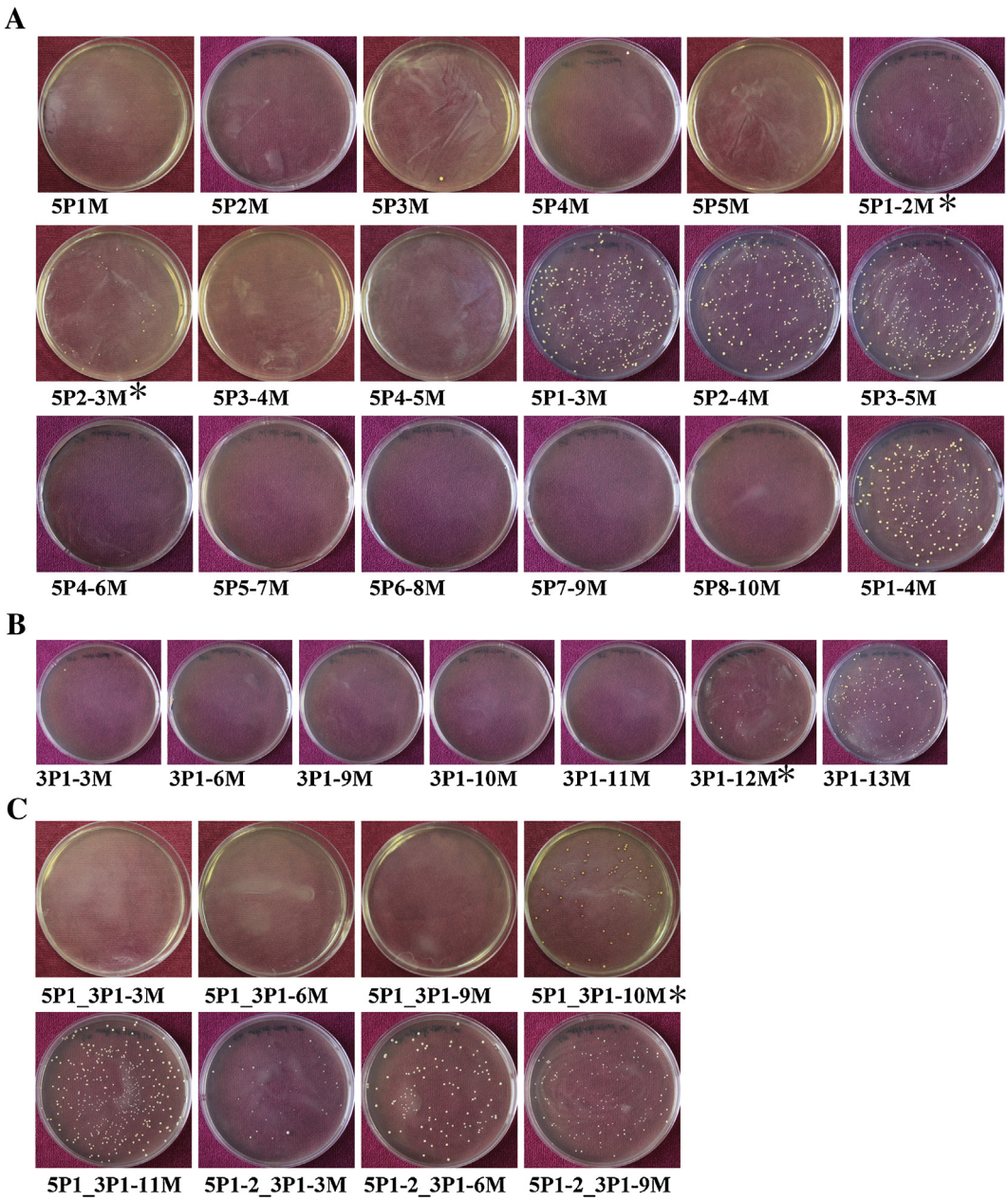

Fig. 6. Base pairing between crRNA and target in the 5′ end is more important than that in the 3′end. The transformation results of *S. aureus* AH1 WT strain with stepwise mutant plasmids 5PM (A), 3PM (B), and 5P_3PM (C). The mutant plasmids were constructed based on the perfectly matched plasmid pRproS29I, and the introduced mutations in target region were used to indicate each transformation experiment. The asterisks were used to mark the plates where the colonies were smaller. Plasmids used in these experiments were extracted from *S. aureus* AH1 CRISPR knockout strain to ensure high experimental sensitivity. Three repeat experiments were performed to confirm the results, and the representative pictures were shown.

et al., 2014). The significance of CRISPR-Cas systems in *S. aureus* should not be undervalued since they were retained in some clinical strains. And the relevance between staphylococcal CRISPR element and pathogenesis also deserves further investigation.

The CRISPR array contains conserved repeats and distinct spacers. The recombination between repeats can result in repeat-spacer unit loss (Jiang et al., 2013), and high identity between repeats may facilitate this process. Therefore, we suspected that sequence variation in the trailer-end repeat may have been selected to avoid the complete loss of the whole CRISPR array. Possibly, the single repeat found in *S. aureus* AH1 is the remnant of a lost CRISPR array. The CRISPR spacers mostly originate from invading elements. However, only a few of the identified spacers in this study were found to have complete homologous sequences in staphylococcal phages or MGEs (Supplementary Table S3 in the online version at DOI: 10.1016/j.ijmm.2016.08.005), possibly because a large number of staphylococcal phages have not been sequenced. In addition, partially complementary sequences of spacers were found on phages, which are likely to be the homologous sequence of their protospacers and could be targeted by CRISPR-Cas immunity. Most target sequences identified were located in the lytic region of the prophage, where the genes are not expressed under normal conditions. A typical example was found in *S. aureus* SH3, where two identical partially complementary sequences of S24 with only five mismatches were located in prophage region. The co-existence of the spacer and its target can be explained with the conditional tolerance of temper-

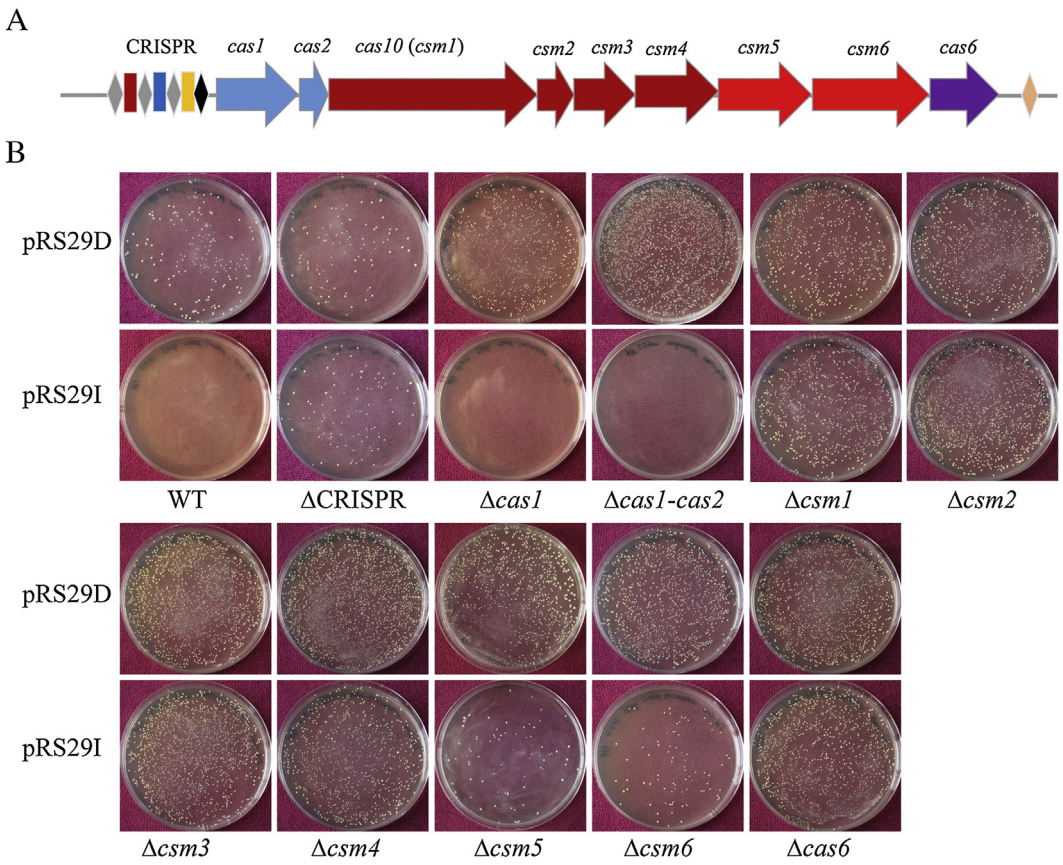

Fig. 7. The Csm complex and Cas6 are required for antiplasmid CRISPR-Cas immunity. (A) The CRISPR-Cas system in *S. aureus* AH1. The repeats (diamond) and spacers (rectangle) were coloured, and the single repeat downstream of the *cas* operon was also indicated. The nine *cas*/*csm* genes were coloured to indicate their involvement in CRISPR-Cas immunity according to the results from this study. (B) The transformation results of *S. aureus* AH1 *cas*/*csm* deletion strains with S29 target plasmids. Plasmids used in these experiments were extracted from *S. aureus* AH1 CRISPR knockout strain to ensure high experimental sensitivity. Three repeat experiments were carried out to confirm the results, and the representative pictures were shown.

ate phages (Goldberg et al., 2014). This also suggests that some CRISPR-Cas systems can benefit their hosts by helping to maintain the lysogenic state of prophages.

The base paring between crRNA and its target is necessary for CRISPR-Cas immunity, but mismatching can also be tolerated. In this study, the first five nucleotides in the 5′ end seem to be more important, and the seed sequence in type III-A CRISPR-Cas system, if exists, is supposed to be located in 5′-proximal region. Mismatching of about 12 nucleotides in the 3′ end did not impair the immunity, which is in agreement with the finding that secondary processing of the crRNA results in 6-nt interval trimming, and the shortest functional crRNA was 12-nt trimmed (Hatoum-Aslan et al., 2013). The spacer sequence itself is also important in CRISPR-Cas immunity, as influence of mismatches was reported to be sequence-dependent (Hsu et al., 2013; Maniv et al., 2016). Thus, the important sites for other spacers may not be exactly the same as indicated in this study. The difference may result from the GC content of spacer sequence and the secondary structure of crRNA.

Both *S. aureus* AH1 and *S. epidermidis* RP62a have a CRISPR array with three functional spacers. In this study, Csm1, Csm2, Csm3, Csm4 and Cas6 were essential for type III-A CRISPR-Cas immunity, but Cas1 and Cas2 were not involved in interference. These data are consistent with the previous reported Cas functions in *S. epidermidis* RP62a (Hatoum-Aslan et al., 2014). Whereas, Csm5 and Csm6 seemed not essential for preventing plasmid transformation, which is not in agreement with the previous report, in which Csm5 and Csm6 were required for plasmid conjugation (Hatoum-Aslan et al., 2014).

In conclusion, the clinical isolates of *S. aureus* with type III-A CRISPR-Cas systems identified in this study, together with the sequence analyses and functional studies should provide new insights into further functional and mechanistic understanding of CRISPR-Cas system in *S. aureus*.

4. Material and methods

4.1. Bacterial strains and culture conditions

The sources of the clinical isolates were listed in Supplementary Table S1 in the online version at DOI: 10.1016/j.ijmm.2016.08.005. *Escherichia coli* and *S. aureus* strains were cultivated with shaking (220 rpm) at 37 °C, in lysogeny broth (LB) medium (Oxoid) and tryptic soy broth (TSB) medium (Difco), respectively, or on corresponding solid medium supplemented with agar. When required, the media were supplemented with 15 μg/ml chloromycetin for *S. aureus* and 100 μg/ml ampicillin for *E. coli*. The *S. aureus* cultures for genome or plasmid extraction were pre-digested with a digestion buffer containing 40 U/mL lysostaphin, 10 mg/ml lysozyme, and 10% (v/v) glycerol. Genomic DNA was prepared using a standard protocol for Gram-positive bacteria (Flamm et al., 1984).

4.2. Whole-genome sequencing, annotation, and bioinformatics analysis

Whole-genome shotgun sequencing was performed on an Illumina HiSeq 2500 platform with paired-end libraries (read length,

100 nt). For each genome, more than 20 M high-quality reads (Q20 > 95%, Q30 > 85%) were obtained after filtering. Reads were subjected to *de novo* assembly using SPAdes software (Nurk et al., 2013). Then, low-coverage contigs (coverage <25×) and short contigs (length <500 bp) were removed. Contigs were subject to gene prediction and annotation using the MicroScope platform (Vallenet et al., 2013). The sequences and annotations were deposited in the European Nucleotide Archive (ENA).

Multilocus sequence typing (MLST) of these isolates was conducted in the MLST webserver (http://www.mlst.net/) (Enright et al., 2000). The staphylococcal protein A (*spa*) typing was conducted based on analysis of the short sequence repeats of *spa* (Mellmann et al., 2007). The SCC*mec* typing was determined following guidelines (International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (IWG-SCC), 2009). Besides, the putative prophages were predicted, annotated and visualized using PHAST (Zhou et al., 2011). CRISPRs were predicted using CRISPR finder (Grissa et al., 2007), Piler-CR (Edgar, 2007), and/or CRT (Bland et al., 2007), and the results were combined and manually curated. To analyse pre-existing staphylococcal CRISPR-Cas systems, publically available staphylococcal genome sequences were retrieved from GenBank (April 2016) and subjected to similar analyses. Multiple sequence alignments were performed using MEGA (Tamura et al., 2013), and spacer homologues were determined using the BLAST+ software (blastn-short task was employed). Because the lengths of the spacers were different, the BLAST scores of the spacer homologues were normalized through dividing by the corresponding spacer length and multiplying by 100 (normalized BLAST score = original BLAST score * 100/spacer length).

4.3. Plasmids construction

Sequences of spacers S29, S30, and S31 were ligated into plasmids pRMC2 or pALC-GFP by SLICE (Zhang et al., 2012), each in two orientations. Briefly, fragments containing the spacers were designed to have overlaps with the corresponding plasmids, and the fragments were first obtained by annealing two overlapping primers, and then ligated to linearized plasmids by SLICE with the recommended procedures and conditions as previously described (Zhang et al., 2015). With the same method, the supposed target sequences of S29 and S30 in phages were also ligated into pRMC2 to get the plasmid pRS29I-ph66,-phGRCS,-phPsa3, and pRS30I-phvB. To construct pRproS29I, to which S29 perfectly matches, a fragment was amplified by PCR using primer pair P34 and P38 with pRS29I-ph66 as a template, and the fragment was then digested with EcoRI/XbaI and cloned into pRMC2. The mutant plasmids were constructed based on pRproS29I. In general, mutations were introduced by two consecutive PCR steps with internal primers carrying mutations, followed by an overlap PCR with the external primer pair P37 and P38, and the fragments containing mutations were then digested with EcoRI/XbaI and cloned into pRMC2. To create CRISPR knockout strain and in-frame deletions of the *cas*/*csm* genes in *S. aureus* AH1, the homologous recombinant plasmid pBTs was employed as previously described (Hu et al., 2015). All the primers used in this study were listed in Supplementary Table S2 in the online version at DOI: 10.1016/j.ijmm.2016.08.005.

4.4. Preparation of S. aureus electrocompetent cells and transformation experiments

S. aureus cells were streaked and grown overnight on TSB agar, a single colony was picked and inoculated in 2 ml of TSB. After 10 h growth (220 rpm), the culture was diluted in fresh medium without antibiotic to an optical density at 600 nm (OD_{600}) of 0.05, and continued to grow about 2.5 h to OD_{600} of 0.5-0.7. The cells were cooled down on ice for 5 min before collected by centrifugation (5,000*g* for 10 min) at 4 °C. Two or three washes were performed using chilled, sterile sucrose (0.5 M), and cells were ultimately resuspended in 1/50 vol of sucrose (0.5 M) and 100 ul of aliquots were prepared for storage at −80 °C.

Plasmids for detection of the CRISPR-Cas immunity were extracted from *S. aureus* RN4220 or AH1 CRISPR knockout strain. The concentrations of plasmids were measured using a NanoDrop2000c Spectrophotometer (Thermo Scientific) and adjusted to 100 ng/μl. Aliquots (100 ul) of competent cells were transformed with 500 ng of plasmid DNA per transformation using a MicroPulser electroporator (Bio-Rad) following the parameters: 2.5 kV, 50 μF, 200 Ω, 2 mm (the preset program Ec2). After electroporation, cells were immediately resuspended in TSB to a final volume of 500 ml and recovered at 37 °C for 1 h with shaking (220 rpm) before plated on solid medium containing chloromycetin (15 μg/ml). Plates were incubated at 37 °C for 18–24 h, or 36–48 h for the slower grown colonies. For each experiment, plasmids from two colonies were tested with at least two repeats to confirm the results, and the representative pictures were shown.

Competing financial interests

The authors declare no competing financial interests.

Nucleotide sequence accession numbers

Samples information, short sequencing reads, and their annotations were deposited in the European Nucleotide Archive (ENA), with the following numbers: PRJEB8895, PRJEB8896, PRJEB8897, PRJEB8898, PRJEB8899, PRJEB8900.

Funding

This work was supported by the National Natural Science Foundation of China [31670133]

Acknowledgements

The authors thank the LABGeM and the National Infrastructure France Genomique for helping with the genome annotation and data submission processing. We also thank Bai G. for her technical assistance.

References

Ballhausen, B., Jung, P., Kriegeskorte, A., Makgotlho, P.E., Ruffing, U., von Müller, L., Köck, R., Peters, G., Herrmann, M., Ziebuhr, W., Becker, K., Bischoff, M., 2014. LA-MRSA CC398 differ from classical community acquired-MRSA and hospital acquired-MRSA lineages: functional analysis of infection and colonization processes. Int. J. Med. Microbiol. 304, 777–786, http://dx.doi.org/10.1016/j.ijmm.2014.06.006.

Bland, C., Ramsey, T.L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N.C., Hugenholtz, P., 2007. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinf. 8, 209, http://dx.doi.org/10.1186/1471-2105-8-209.

Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiol. Read. Engl. 151, 2551–2561, http://dx.doi.org/10.1099/mic.0.28048-0.

Bondy-Denomy, J., Davidson, A.R., 2014. To acquire or resist: the complex biological effects of CRISPR-Cas systems. Trends Microbiol. 22, 218–225, http://dx.doi.org/10.1016/j.tim.2014.01.007.

Carte, J., Wang, R., Li, H., Terns, R.M., Terns, M.P., 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. 22, 3489–3496, http://dx.doi.org/10.1101/gad.1742908.

Chao, G., Bao, G., Jiao, X., 2014. Molecular epidemiological characteristics and clonal genetic diversity of *Staphylococcus aureus* with different origins in China. Foodborne Pathog. Dis. 11, 503–510, http://dx.doi.org/10.1089/fpd.2013.1717.

Corrigan, R.M., Foster, T.J., 2009. An improved tetracycline-inducible expression vector for *Staphylococcus aureus*. Plasmid 61, 126–129, http://dx.doi.org/10.1016/j.plasmid.2008.10.001.

Crooks, G.E., Hon, G., Chandonia, J.-M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190, http://dx.doi.org/10.1101/gr.849004.

Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., Semenova, E., 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat. Commun. 3, 945, http://dx.doi.org/10.1038/ncomms1937.

Edgar, R.C., 2007. PILER-CR: fast and accurate identification of CRISPR repeats. BMC Bioinf. 8, 18, http://dx.doi.org/10.1186/1471-2105-8-18.

Enright, M.C., Day, N.P., Davies, C.E., Peacock, S.J., Spratt, B.G., 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J. Clin. Microbiol. 38, 1008–1015.

Flamm, R.K., Hinrichs, D.J., Thomashow, M.F., 1984. Introduction of pAM beta 1 into *Listeria monocytogenes* by conjugation and homology between native *L. monocytogenes* plasmids. Infect. Immun. 44, 157–161.

Goldberg, G.W., Jiang, W., Bikard, D., Marraffini, L.A., 2014. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature 514, 633–637, http://dx.doi.org/10.1038/nature13637.

Golding, G.R., Bryden, L., Levett, P.N., McDonald, R.R., Wong, A., Wylie, J., Graham, M.R., Tyler, S., Van Domselaar, G., Simor, A.E., Gravel, D., Mulvey, M.R., 2010. Livestock-associated methicillin-resistant *Staphylococcus aureus* sequence type 398 in humans, Canda. Emerg. Infect. Dis. 16, 587–594, http://dx.doi.org/10.3201/eid1604.091435.

Golding, G.R., Bryden, L., Levett, P.N., McDonald, R.R., Wong, A., Graham, M.R., Tyler, S., Van Domselaar, G., Mabon, P., Kent, H., Butaye, P., Smith, T.C., Kadlec, K., Schwarz, S., Weese, S.J., Mulvey, M.R., 2012. Whole-genome sequence of livestock-associated st398 methicillin-resistant *Staphylococcus aureus* Isolated from Humans in Canada. J. Bacteriol. 194, 6627–6628, http://dx.doi.org/10.1128/JB.01680-12.

Grissa, I., Vergnaud, G., Pourcel, C., 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinf. 8, 172, http://dx.doi.org/10.1186/1471-2105-8-172.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., Terns, M.P., 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 139, 945–956, http://dx.doi.org/10.1016/j.cell.2009.07.040.

Hatoum-Aslan, A., Maniv, I., Marraffini, L.A., 2011. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proc. Natl. Acad. Sci. U. S. A. 108, 21218–21222, http://dx.doi.org/10.1073/pnas.1112832108.

Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W., Marraffini, L.A., 2013. A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. J. Biol. Chem. 288, 27888–27897, http://dx.doi.org/10.1074/jbc.M113.499244.

Hatoum-Aslan, A., Maniv, I., Samai, P., Marraffini, L.A., 2014. Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. J. Bacteriol. 196, 310–317, http://dx.doi.org/10.1128/JB.01130-13.

Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., Marraffini, L.A., 2015. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature 519, 199–202, http://dx.doi.org/10.1038/nature14245.

Holt, D.C., Holden, M.T.G., Tong, S.Y.C., Castillo-Ramirez, S., Clarke, L., Quail, M.A., Currie, B.J., Parkhill, J., Bentley, S.D., Feil, E.J., Giffard, P.M., 2011. A very early-branching *Staphylococcus aureus* lineage lacking the larotenoid pigment staphyloxanthin. Genome Biol. Evol. 3, 881–895, http://dx.doi.org/10.1093/gbe/evr078.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31, 827–832, http://dx.doi.org/10.1038/nbt.2647.

Hu, J., Zhang, X., Liu, X., Chen, C., Sun, B., 2015. Mechanism of reduced vancomycin susceptibility conferred by walK mutation in community-acquired methicillin-resistant *Staphylococcus aureus* strain MW2. Antimicrob. Agents Chemother. 59, 1352–1355, http://dx.doi.org/10.1128/AAC.04290-14.

International working group on the classification of staphylococcal cassette chromosome elements (IWG-SCC), 2009. Classification of staphylococcal cassette chromosome mec (SCCmec): Guidelines for reporting novel SCCmec elements. Antimicrob. Agents Chemother. 53, 4961–4967, http://dx.doi.org/10.1128/AAC.00579-09.

Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B.R., Marraffini, L.A., 2013. Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. PLoS Genet. 9, e1003844, http://dx.doi.org/10.1371/journal.pgen.1003844.

Jiang, W., Samai, P., Marraffini, L.A., 2016. Degradation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity. Cell 164, 710–721, http://dx.doi.org/10.1016/j.cell.2015.12.053.

Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č., Siksnys, V., 2016. Spatiotemporal control of type III-A CRISPR-Cas immunity: coupling DNA degradation with the target RNA recognition. Mol. Cell 62, 295–306, http://dx.doi.org/10.1016/j.molcel.2016.03.024.

Kinnevey, P.M., Shore, A.C., Brennan, G.I., Sullivan, D.J., Ehricht, R., Monecke, S., Slickers, P., Coleman, D.C., 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome mec (SCCmec)-SCC-SCCCRISPR composite element in Irish hospitals. Antimicrob. Agents Chemother. 57, 524–531, http://dx.doi.org/10.1128/aac.01689-12.

Li, M., Wang, R., Xiang, H., 2014. Haloarcula hispanica CRISPR authenticates PAM of a target sequence to prime discriminative adaptation. Nucleic Acids Res. 42, 7226–7235, http://dx.doi.org/10.1093/nar/gku389.

Lindsay, J.A., 2014. *Staphylococcus aureus* genomics and the impact of horizontal gene transfer. Int. J. Med. Microbiol. 304, 103–109, http://dx.doi.org/10.1016/j.ijmm.2013.11.010.

Lowy, F.D., 1998. *Staphylococcus aureus* infections. N. Engl. J. Med. 339, 520–532, http://dx.doi.org/10.1056/NEJM199808203390806.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Terns, R.M., Terns, M.P., White, M.F., Yakunin, A.F., Garrett, R.A., van der Oost, J., Backofen, R., Koonin, E.V., 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. 13, 722–736, http://dx.doi.org/10.1038/nrmicro3569.

Maniv, I., Jiang, W., Bikard, D., Marraffini, L.A., 2016. Impact of different target sequences on type III CRISPR-Cas immunity. J. Bacteriol. 198, 941–950, http://dx.doi.org/10.1128/JB.00897-15.

Marraffini, L.A., Sontheimer, E.J., 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322, 1843–1845, http://dx.doi.org/10.1126/science.1165771.

Marraffini, L.A., Sontheimer, E.J., 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 463, 568–571, http://dx.doi.org/10.1038/nature08703.

Marraffini, L.A., 2015. CRISPR-Cas immunity in prokaryotes. Nature 526, 55–61, http://dx.doi.org/10.1038/nature15386.

Mellmann, A., Weniger, T., Berssenbrugge, C., Rothganger, J., Sammeth, M., Stoye, J., Harmsen, D., 2007. Based upon repeat pattern (BURP): an algorithm to characterize the long-term evolution of *Staphylococcus aureus* populations based on spa polymorphisms. BMC Microbiol. 7, 98, http://dx.doi.org/10.1186/1471-2180-7-98.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., Soria, E., 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182, http://dx.doi.org/10.1007/s00239-004-0046-3.

Niewoehner, O., Jinek, M., 2016. Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA N. Y. N 22, 318–329, http://dx.doi.org/10.1261/rna.054098.115.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A.A., Korobeynikov, A., Lapidus, A., Prjibelski, A.D., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., Clingenpeel, S.R., Woyke, T., McLean, J.S., Lasken, R., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J. Comput. Biol. J. Comput. Mol. Cell Biol. 20, 714–737, http://dx.doi.org/10.1089/cmb.2013.0084.

Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., Marraffini, L.A., 2015. Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. Cell 161, 1164–1174, http://dx.doi.org/10.1016/j.cell.2015.04.027.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Res. 39, 9275–9282, http://dx.doi.org/10.1093/nar/gkr606.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J.J., Severinov, K., 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. U. S. A. 108, 10098–10103, http://dx.doi.org/10.1073/pnas.1104144108.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., Doudna, J.A., 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507, 62–67, http://dx.doi.org/10.1038/nature13011.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729, http://dx.doi.org/10.1093/molbev/mst197.

Vallenet, D., Belda, E., Calteau, A., Cruveiller, S., Engelen, S., Lajus, A., Le Fevre, F., Longin, C., Mornico, D., Roche, D., Rouy, Z., Salvignol, G., Scarpelli, C., Thil Smith, A.A., Weiman, M., Medigue, C., 2013. MicroScope–an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. Nucleic Acids Res. 41, D636–D647, http://dx.doi.org/10.1093/nar/gks1194.

Vestergaard, G., Garrett, R.A., Shah, S.A., 2014. CRISPR adaptive immune systems of archaea. RNA Biol. 11, 156–167, http://dx.doi.org/10.4161/rna.27990.

Wakefield, N., Rajan, R., Sontheimer, E.J., 2015. Primary processing of CRISPR RNA by the endonuclease Cas6 in *Staphylococcus epidermidis*. FEBS Lett. 589, 3197–3204, http://dx.doi.org/10.1016/j.febslet.2015.09.005.

Wang, Y., Hu, M., Liu, Q., Qin, J., Dai, Y., He, L., Li, T., Zheng, B., Zhou, F., Yu, K., Fang, J., Liu, X., Otto, M., Li, M., 2016. Role of the ESAT-6 secretion system in virulence of the emerging community-associated *Staphylococcus aureus* lineage ST398. Sci. Rep. 6, 25163, http://dx.doi.org/10.1038/srep25163.

Westra, E.R., Swarts, D.C., Staals, R.H.J., Jore, M.M., Brouns, S.J.J., van der Oost, J., 2012. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu. Rev. Genet. 46, 311–339, http://dx.doi.org/10.1146/annurev-genet-110711-155447.

Westra, E.R., Buckling, A., Fineran, P.C., 2014. CRISPR-Cas systems: beyond adaptive immunity. Nat. Rev. Microbiol. 12, 317–326, http://dx.doi.org/10.1038/nrmicro3241.

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S.M., Ma, W., Doudna, J.A., 2009. Structural basis for DNase activity of a conserved protein implicated in

CRISPR-mediated genome defense. Structure 17, 904–912, http://dx.doi.org/10.1016/j.str.2009.03.019.
Wiedenheft, B., van Duijn, E., Bultema, J.B., Bultema, J., Waghmare, S.P., Waghmare, S., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J.R., Heck, A., Boekema, E.J., Boekema, E., Dickman, M.J., Dickman, M., Doudna, J.A., 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. U. S. A. 108, 10092–10097, http://dx.doi.org/10.1073/pnas.1102716108.
Zebec, Z., Manica, A., Zhang, J., White, M.F., Schleper, C., 2014. CRISPR-mediated targeted mRNA degradation in the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res. 42, 5280–5288, http://dx.doi.org/10.1093/nar/gku161.
Zhang, Y., Werling, U., Edelmann, W., 2012. SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 40, e55, http://dx.doi.org/10.1093/nar/gkr1288.
Zhang, S., Ma, R., Liu, X., Zhang, X., Sun, B., 2015. Modulation of ccrAB expression and SCCmec excision by an inverted repeat element and SarS in methicillin-resistant *Staphylococcus aureus*. Antimicrob. Agents Chemother. 59, 6223–6232, http://dx.doi.org/10.1128/AAC.01041-15.
Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S., 2011. PHAST: a fast phage search tool. Nucleic Acids Res. 39, W347–352, http://dx.doi.org/10.1093/nar/gkr485.
van der Oost, J., Westra, E.R., Jackson, R.N., Wiedenheft, B., 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. 12, 479–492, http://dx.doi.org/10.1038/nrmicro3279.