Curr. Issues Mol. Biol. (2019) 32: 1-38. DOI: https://dx.doi.org/10.21775/cimb.032.001

CRISPR-Cas Systems in Streptococci

Tao Gong[#], Miao Lu[#], Xuedong Zhou, Anqi Zhang, Boyu Tang, Jiamin Chen, Meiling Jing, Yuqing Li*

State Key Laboratory of Oral Diseases, National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China

Tao Gong[#] and Miao Lu[#] contributed equally to this work.

*Corresponding author: liyuqing@scu.edu.cn

Abstract

Streptococci are one of the most important and common constituents of the host's microbiota and can colonize and live in the upper respiratory and urogenital tract of humans and animals. The CRISPR-Cas systems (i.e., clustered regularly interspaced short palindromic repeat, with CRISPR-associated proteins) found in bacteria and archaea provide sequence-based adaptive immunity against mobile genetic elements, especially in the streptococci. Here, recent research progress on CRISPR-Cas systems in the streptococci is reviewed, including their classification (mainly type I, type II, and type III), physiological function, defense mechanism (CRISPR adaptation, crRNA biogenesis, and target interference) and applications, which are useful for a better understanding of the functions of such systems. Finally, the advances

that have been made in streptococci may help in the discovery of further novel CRISPR-Cas systems for use in new technologies and applications in other species.

1. Introduction

Streptococci, a heterogeneous group of Gram-positive facultative anaerobic bacteria, commonly colonize and live in the upper respiratory and urogenital tract of humans and animals, but also are present in other hosts. The genus *Streptococcus* consists of more than 100 diverse species (Willenborg and Goethe, 2016) that tend to be adapted to their hosts, reflecting co-evolution. Thus, streptococci are very important components of the host's microbiota. Although streptococci are commonly regarded as commensals, streptococcal diseases range from moderately serious infections such as impetigo, cellulitis or pharyngitis to severe life-threatening diseases such as meningitis, pneumonia, endocarditis and septicemia. Of the streptococci only *Streptococcus thermophilus* does not have pathogenicity-related genes and is therefore regarded as nonpathogenic (Hols et al., 2005). However, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* and *Streptococcus agalactiae* are considered as important human pathogens. For instance, *S. pyogenes* usually colonizes skin epithelial surfaces or the throat and can cause a series of clinical diseases such as necrotizing fasciitis or streptococcal toxic shock syndrome in humans (Hamada et al., 2015); *S. pneumoniae* is the most common cause of community-acquired pneumonia and meningitis (Lynch and Zhanel, 2009); *S. mutans* is regarded as the major etiological agent of dental caries (Tanzer et al., 2001) and *S. agalactiae* is an

emerging cause of septic arthritis in some countries (Louthrenoo et al., 2014). Nevertheless, other streptococcal species related to the human microbiota can also cause invasive or noninvasive diseases, including *S. sanguinis* (endocarditis) (Jung et al., 2012) and *S. gordonii* (septic arthritis) (Yombi et al., 2012).

Streptococcus species undergo frequent and extensive horizontal gene transfer (HGT), through transformation, transduction, or conjugation (Choi et al., 2012). HGT has played a major role in genomic evolution, helping in the acquisition of new traits that impact processes such as metabolic pathways, pathogenesis and antibiotic resistance. For example, prophages make up to 14% of the *S. pyogenes* genome and encode virulence factors including DNases, pyrogenic exotoxins and the phospholipase SlaA (Fischetti, 2007). However, some innate strategies have also evolved to protect against invading foreign genetic material, including ComX or SigX factors, CSP (competence stimulating peptide), restriction-modification (R-M) systems or recombination machinery (Fontaine et al., 2015).

Besides the innate immune strategies mentioned above, some bacteria also possess an alternative adaptive immune system. CRISPR-Cas systems (clustered regularly interspaced short palindromic repeat, with CRISPR-associated proteins) are found in 40% of bacteria and 90% of archaea and provide sequence-based adaptive immunity against mobile genetic elements such as phages, invasive conjugative plasmids and transposable elements (Marraffini, 2015). The two essential components, CRISPR loci and

their associated *cas* genes, constitute a small RNA-guided system that helps prokaryotes ward off invasive genetic elements. A CRISPR locus contains arrays of 'spacers' derived from the genomes of previous invaders, usually viruses or plasmids, and those critical 'spacer' sequences are inserted between the repeat elements of natural CRISPR loci. A cluster of *cas* genes is located upstream of a CRISPR locus and different CRISPR-Cas systems possess diverse *cas* genes (Makarova et al., 2011). CRISPR-Cas systems were first identified as a specific sequence-based barrier against three major routes of foreign DNA introduction: plasmid conjugation, phage transduction and DNA transformation (Marraffini, 2015). To date, the CIRSPR-Cas system is the only known prokaryotic immune system that enables precise recognition and targeting of phages once a specific sequence of from its genome has been 'memorized'. Importantly, this 'memory' can be passed down to subsequent generations.

In this review, we discuss current knowledge on the distribution, physiological functions and mechanisms of CRISPR-Cas systems encoded in streptococci, their role in the biology of these organisms and their applications.

2. Distribution of CRISPR-Cas systems in streptococci

CRISPR-Cas systems are divided into two distinct classes, characterized by their effector module design (Amitai and Sorek, 2016). Class 1 CRISPR-Cas systems, containing the most common and diverse type I systems, type III predominantly in archaea, as well as type IV systems, are defined by the presence of a multi-subunit crRNA-effector complex. Class 2 CRISPR-Cas systems, including the most prevalent type II, type V, and a new type VI, are

defined by the presence of a large, single, multi-domain crRNA-effector module. Each type can be further partitioned into several subtypes, up to 19 subtypes in all (Koonin et al., 2017). More and more CRISPR-Cas systems are being revealed in the streptococci. As shown in Table 1, streptococci mainly harbor three types of CRISPR-Cas systems: type I, type II and type III, in addition to the individual type V and unknown CRISPR loci. Of these, the number of type II systems, especially subtype II-A, is much larger than the other types, suggesting that streptococci mainly depend on type II systems to defend against invading foreign genetic material or participate in physiology and pathogenesis, whereas many types remain to be identified. The CRISPR loci of several representative types are shown in Figure 1.

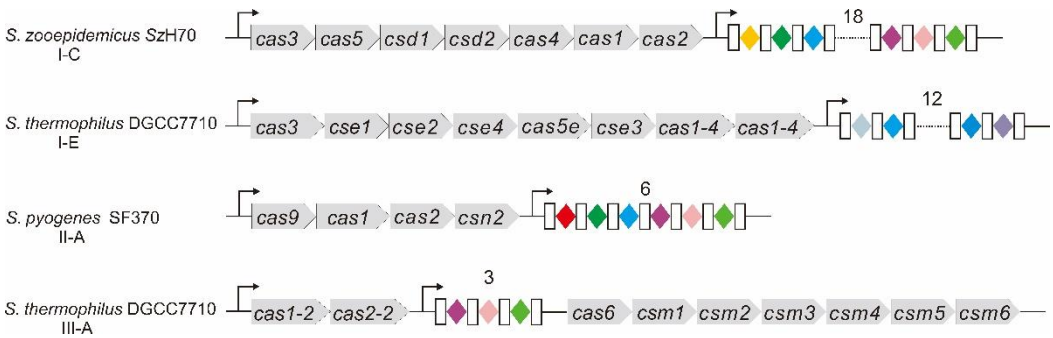

Figure 1. The characteristics of four representative CRISPR-Cas systems in streptococci.

Table 1 Types of streptococcal CRISPR-Cas systems covered in this review.

Species	Type I			Type II			Type III		Type V	Unknown	References
	C	E	Unknown	A	C	Unknown	A	Unknown	A		
S. agalactiae	■		■	■	■		■				Lier et al., 2015; Lopez-Sanchez et al., 2012
S. anginosus				■	■				■		Olson et al., 2013; Makarova et al., 2015
S. australis										■	Maeda et al., 2011
S. constellatus subsp.*constellatus*	■			■							Olson et al., 2013
S. dysgalactiae subsp. *equisimilis*	■		■	■		■					Shimomura et al., 2011; Makarova et al., 2015
S. gallolyticus				■			■				Louwen et al., 2014
S. gordonii				■	■		■				Louwen et al., 2014; Makarova et al., 2015
S. infantarius subsp. *infantarius*				■						■	Jans et al., 2013; Makarova et al., 2015
S. iniae				■						■	Zhang et al., 2014; Makarova et al., 2015
S. intermedius				■							Olson et al., 2013
S. lutetiensis				■							Makarova et al., 2015
S. macedonicus				■							Makarova et al., 2015
S. mitis										■	Kilian et al., 2014
S. mutans	■	■		■			■				Serbanescu et al., 2015; Louwen et al., 2014
S. oligofermentans	■						■				Makarova et al., 2015
S. oralis										■	Kilian et al., 2014
S. parasanguinis	■									■	Makarova et al., 2015
S. parauberis										■	Nho et al., 2013
S. pasteurianus				■						■	Makarova et al., 2015
S. phocae subsp. *phocae*						■					Bethke and Avendano-Herrera, 2017
S. phocae subsp. *salmonis*						■					Bethke and Avendano-Herrera, 2017
S. pyogenes	■			■			■				Hynes et al., 2017; Louwen et al., 2014
S. ruminantium										■	Tohya et al., 2018
S. salivarius		■		■							Makarova et al., 2015
S. sanguinis	■	■		■			■				Louwen et al., 2014
S. suis				■						■	Yao et al., 2015; Makarova et al., 2015
S. thermophilus		■		■			■				Makarova et al., 2015
S. troglodytae										■	Okamoto et al., 2016
S. uberis						■		■			Hossain et al., 2015
S. zooepidemicus	■			■							Waller and Robinson, 2013; Makarova et al., 2015

2.1 Type I CRISPR-Cas systems

As shown in Table 1, 30% (9/30) of analyzed streptococci harbor type I-C, 13.3% (4/30) have type I-E and there are two unknown CRISPR loci (2/30). All type I loci contain the signature gene *cas3* (or its variant *cas3´*), which encodes a helicase protein with a demonstrated capacity to unwind double-stranded DNA (dsDNA) and RNA-DNA duplexes (Huo et al., 2014). Type I systems are currently divided into seven sub-types, I-A to I-F and I-U (Makarova et al., 2011), of which subtype I-B, containing the complete gene locus (*cas1-cas2-cas3-cas4-cas5-cas6-cas7-cas8*) seems to be best defined by ancestral type I gene composition, whereas other subtypes are divergent derivatives of type I-B with loss of different genes or rearranged gene order (Makarova et al., 2013). For example, subtype I-C seems to be a derivative of subtype I-B that lacks *cas6*, which seems to be functionally replaced by *cas5*. Subtype I-E appears to lack *cas4*, suggesting that diverse types may undergo evolution in complex environmental conditions.

2.2 Type II CRISPR-Cas systems

63.3%(19/30) of the analyzed streptococci in Table 1 contain type II CRISPR-Cas systems (mostly type II-A and several type II-C), which are currently divided into three subtypes: II-A, II-B, and II-C (Chylinski et al., 2014). The signature gene for type II systems is *cas9,* which encodes a multi-domain protein that combines the functions of the crRNA-effector complex to cleave double-stranded DNA targets (Gasiunas et al., 2012), and also contributes to adaptation (Heler et al., 2015). The type II-A CRISPR-Cas system is regarded as the minimal CRISPR-Cas system and contains a repeat-spacer array, only four

cas genes (*cas9*, *cas1*, *cas2 and csn2*), a partially complementary trans-acting RNA (tracrRNA) and RNase III (Figure 2). Subtype II-B lacks *csn2* but includes *cas4*, and the Cas4 protein belongs to the PD-(D/E)xK family of nucleases that have been shown to possess 5' to 3' single-stranded DNA (ssDNA) exonuclease activity or both directions (Makarova et al., 2015). While Csn2 functions in spacer integration, the role of Cas4 in the CRISPR-Cas system remains to be identified (Barrangou et al., 2007). Subtype II-C loci only have three protein-coding genes: *cas1*, *cas2,* and *cas9* (Chylinski et al., 2013).

2.3 Type III CRISPR-Cas systems

Of the analyzed streptococci, 26.7% (8/30) possess type III-A CRISPR-Cas systems, and an unknown CRISPR locus (1/30). All type III systems contain the signature gene *cas10* and the Cas10 proteins show significant sequence variation among the diverse type III CRISPR-Cas systems. In fact, there are four type III subtypes: III-A to III-D (Makarova et al., 2015). Type III CRISPR loci have more diverse genetic organization than type I systems due to gene duplications or deletions, and domain insertions or fusions. For instance, type III-C and type III-D system lack the well conserved *cas1* and *cas2* genes (Vestergaard et al., 2014). Moreover, subtype III-A loci usually include *cas1*, *cas2* and *cas6* genes, but most type III-B loci lack these genes and therefore depend on other CRISPR-Cas systems present in the same genome. Both subtype III-A and III-B CRISPR-Cas systems have been shown to co-transcriptionally target RNA and DNA (Samai et al., 2015).

2.4 Other putative CRISPRs in streptococci

About 53.3% (16/30) of the analyzed streptococci possess uncharacterized

CRISPR-Cas systems that have been identified bioinformatically (Table 1). With the rapid development of complete genome sequencing technology, bioinformatic tools have been used to identify CRISPR-Cas systems in many streptococci. These tools include CRISPRFinder (Grissa et al., 2007b), CRISPRdetect (Biswas et al., 2016), CRISPROne (http://omics.informatics.indiana.edu/CRISPRone) for detecting CRISPRs, CRISPRTarget (Biswas et al., 2013) and so forth. For example, prediction of CRISPRs in *S. iniae* SF1 (serotype I) was performed using the CRISPR recognition tool (http://crispr.u-psud.fr/crispr/). It has also been reported that *S. anginosus* contains a type V CRISPR-Cas system (Makarova et al., 2015). Moreover, CRISPRs were detected in the genomes of *Streptococcus infantarius* subsp*. infantarius* CJ18 and ATCC BAA-102T using CRISPRfinder and CRISPRdb (Grissa et al., 2007a, b), but CJ18 contains a smaller number of CRISPR spacers, which indicates a reduced defense capability against foreign DNA (Jans et al., 2013). *S. mitis*, *S. oralis, S. parauberis*, *S. trogolodytae* and *S. ubris* have also been found to possess CRISPR sequences via CRISPRfinder, but these have not been characterized (Kilian et al., 2014; Makarova et al., 2015; Okamoto et al., 2016).

3. Physiological functions of CRISPR-Cas systems

Recent studies have shown that CRISPR-Cas systems can possess alternative roles in pathogenesis or host cell physiology, including in biofilm formation, DNA repair, stress response, immune evasion, quorum sensing and interactions between bacteria (Louwen et al., 2014; Sampson and Weiss, 2014) in addition to their classical function in defense against foreign DNA. For example, the number

of CRISPR sites in forty-five *S. mutan*s strains isolated from plaque samples was associated with the clinical manifestations of early childhood caries (ECC). This study found that the biofilm formation and the ability to synthesize EPS by *S. mutans* strains with two CRISPR sites were significantly improved compared to strains without CRISPR loci, suggesting the CRISPR sites might contribute to the cariogenic potential of *S. mutans* (Chen et al., 2017). The CRISPR-Cas system in *Pseudomonas aeruginosa* plays an important role in modulating biofilm formation. Although the detailed mechanism remains unknown, the study showed that when a certain bacteriophage invades *Pseudomonas aeruginosa*, the CRISPR-Cas system can interact with a gene in the chromosomally integrated prophage to inhibit biofilm formation (Cady and O'Toole, 2011; Zegans et al., 2009). In *S. mutans* UA159, the CRISPR-Cas system was also found to be involved in the stress response, DNA repair and gene expression (Serbanescu et al., 2015). Moreover, Cas1, found in all CRISPR systems, and a crRNA array derived from *Escherichia coli* are essential for DNA repair (Babu et al., 2011). Cas2 also influences the ability of *Legionella pneumophila* to survive within amoebae (Gunderson and Cianciotto, 2013). As amoebae are essential for *Legionella pneumophila* to survive in a complex environment (Abu et al., 1998), the function of Cas2 is predicted to be essential for environmental persistence and subsequent transmission to other hosts.

In pathogens, such as the Gram-negative intracellular pathogen *Francisella novicida*, its type II CRISPR-Cas system plays an important role in bypassing detection by the host's recognition receptors. This system targets an endogenous transcript encoding an immunostimulatory bacterial lipoprotein

(BLP), resulting in mRNA degradation and decreased transcript levels. In contrast, in the absence of this CRISPR system, BLP levels increase slowly, which triggers the activation of a Toll-like Receptor 2 (TLR2)-dependent proinflammatory response, leading to destruction of the bacteria during infection (Jones et al., 2012). Based on the diverse roles of CRISPR systems in different organisms, we speculate that different bacteria may have co-opted distinct CRISPR-Cas components for diverse functional outcomes. As the physiological roles of CRISPR-Cas systems in streptococci are still poorly understood, other functions of these systems remain to be identified.

4. Defense mechanism of CRISPR-Cas systems in streptococci

The molecular mechanism of the CRISPR-Cas system mediates immunity in three phases: (1) CRISPR adaptation, (2) crRNA biogenesis and (3) target interference (see Figure 2). The adaptation phase involves the integration of new spacers derived from plasmids or viruses into the CRISPR array. In the second phase, the CRISPR array is transcribed as a precursor CRISPR RNA (pre-crRNA), which is then processed to form mature crRNAs including a part of the repeat and the spacer. During the interference phase, the mature crRNA in combination with single or multiple Cas proteins, functions as a guide to recognize and cleave invading nucleic acids (also known as protospacers), such as plasmids or viruses (Makarova et al., 2015). In some CRISPR-Cas systems (i.e., types I, II and V), a short protospacer adjacent motif (PAM) sequence found adjacent to the targeted protospacer sequence is required for both the adaptation and interference phases. For these CRISPR-Cas systems, the PAM sequence is located on the 3’ or 5’ end of the foreign targeted DNA but is absent

in the CRISPR array, which functions in target recognition between its own and foreign DNA (Leenay and Beisel, 2017). However, CRISPR-Cas systems without PAM sequences possess strategies to avoid self-targeting, such as lacking complementarity between the 5' repeat handle of the crRNA and the 3' flanking region of the target RNA in some type III systems (Tamulaitis et al., 2017).

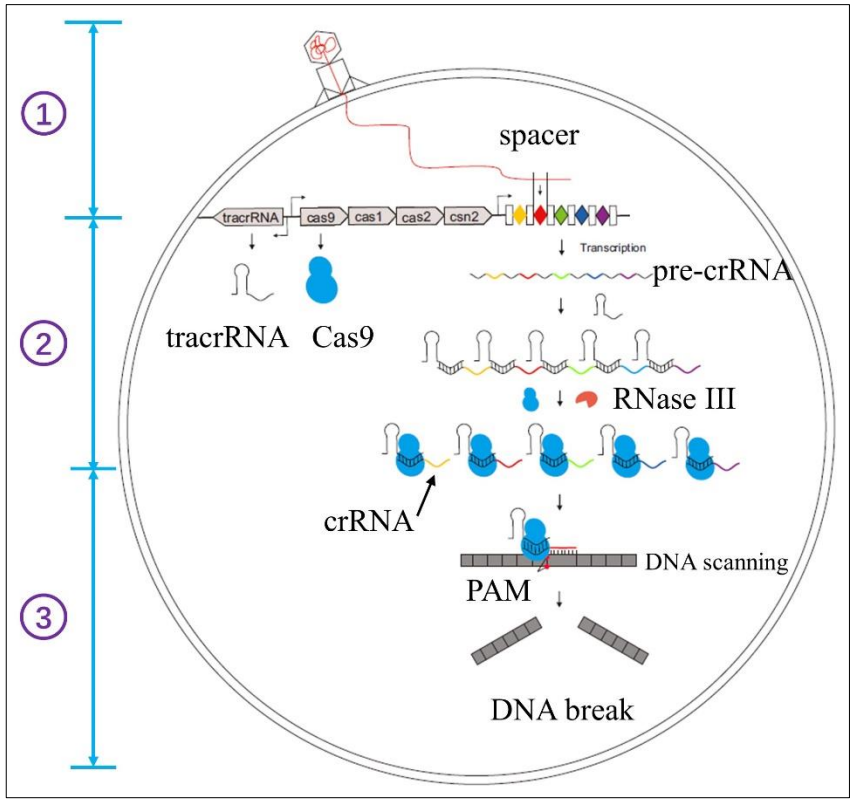

Figure 2. The three stages of CRISPR-Cas9 adaptive immunity: (1) CRISPR adaptation, (2) crRNA biogenesis and (3) target interference.

4.1 CRISPR adaptation

During the CRISPR adaptation phase, short fragments of exogenous DNA derived from plasmids or viruses are incorporated into the CRISPR array within

the bacterial genome, and these protospacers act as new spacer sequences. In this way hosts can combat future attacks by similar invaders and keep a genetic record for similar encounters (Makarova et al., 2015). The adaptation mechanism is best characterized in *E. coli*. The Cas1 and Cas2 proteins, which form a $Cas1_4$-$Cas2_2$ complex (hereafter, Cas1-Cas2), are mainly responsible to spacer integration, whereas some elements such as at least one CRISPR repeat that is part of the leader sequence (Wei et al., 2015) and a few host factors for repair of the insertion sites (Ivancic-Bace et al., 2015) are also required. Spacer acquisition can be divided into three steps: (1) substrate capture, (2) recognition of the CRISPR locus and (3) spacer integration into CRISPR loci (Jackson et al., 2017) (Figure 3).

4.1.1 Cas1-Cas2 substrate capture

In the substrate capture phase, production of prespacers (partially duplexed dsDNA) from foreign DNA can occur by two pathways: naïve CRISPR adaptation and primed CRISPR adaptation (Heler et al., 2014). Cas1/Cas2-mediated acquisition can add a spacer-repeat unit to a CRISPR locus consisting of only one repeat sequence, which is referred to as naïve acquisition. However, primed acquisition occurs when the CRISPR array already possesses spacers matching the target DNA and the presence of spacers can increase the frequency of acquiring another spacer. The ends of the dsDNA prespacer combined with Cas1-Cas2 are separated by tyrosine wedges in each Cas1 dimer, which locks and opens the DNA branch points. Each 3’ end of the prespacer spreads into the active subunits of each corresponding Cas1 dimer (Nunez et al., 2015; Wang et al., 2015) (Figure 3).

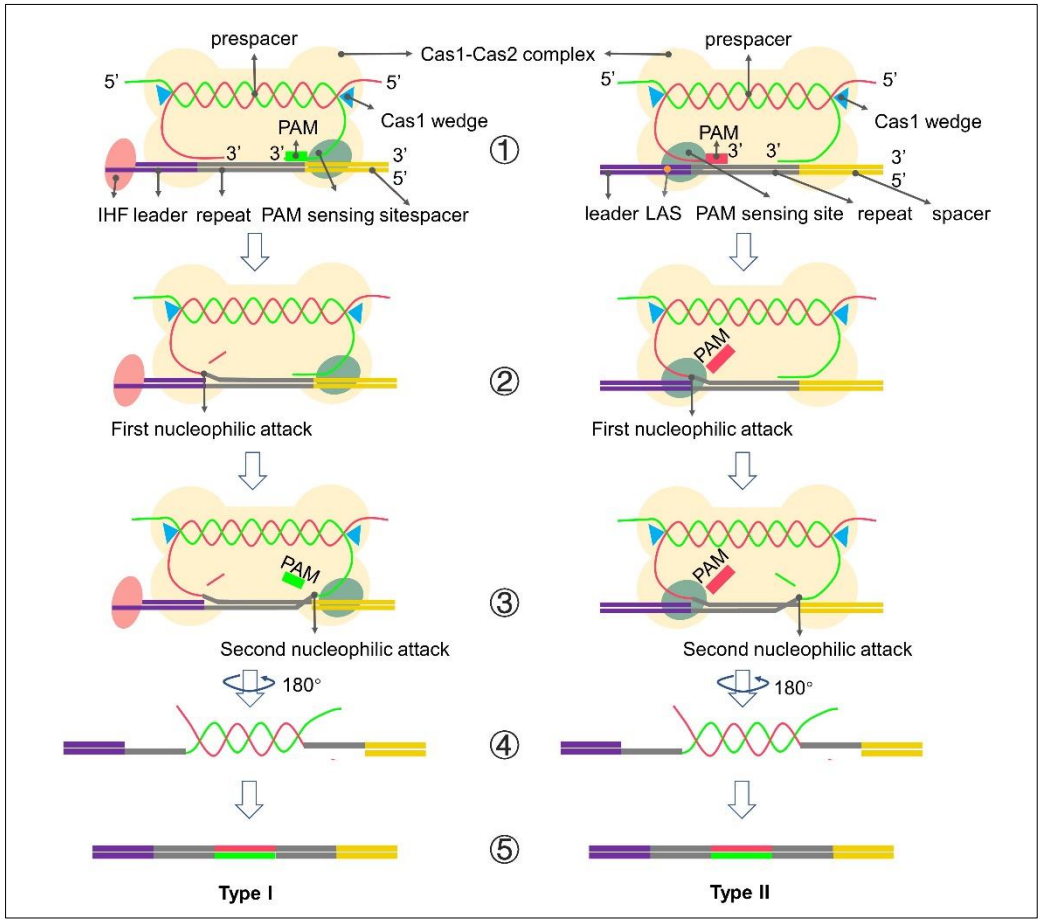

Figure 3. Spacer integration process: (1) Cas1-Cas2 complex load with prespacer into Cas1-Cas2-prespacer complex, which binds leader and first repeat. For type I systems (e.g., I-E in *E. coli*), Cas1-Cas2 binding to the leader-proximal repeat is assisted by integration host factor (IHF), whereas Cas1-Cas2 binding to the leader-proximal repeat in type II systems (e.g., II-A in *S. pyogenes*) need recognize the leader-anchoring site (LAS). (2) The first nucleophilic attack likely occurs at the leader-repeat junction and forms a half-site intermediate. (3) The second nucleophilic attack occurs at the existing repeat-spacer junction into full-site integration. (4) and (5) Host DNA repair mechanisms fill the integration site to acquire perfect integration.

4.1.2 Recognition of the CRISPR array

During recognition, the leader-repeat sequences upstream of the CRISPR arrays directly guide leader-polarized spacer integration and are assisted by direct Cas1-Cas2 complex recognition and host proteins (Wright and Doudna, 2016). For the type I-E system from *E. coli*, leader-repeat recognition is required by the integration host factor (IHF) heterodimer. IHF, which binds the CRISPR leader, helps to localize Cas1-Cas2 to the leader-repeat junction, which also contributes to stabilizing the interaction between Cas1-Cas2 and the leader-repeat, increasing the efficiency of spacer acquisition (Nunez et al., 2016). However, in type II systems from *S. pyogenes*, a short leader-anchoring site (LAS) adjacent to the first repeat has been found to be essential for CRISPR spacer adaptation (Wei et al., 2015) (Figure 3).

4.1.3 Integration into the CRISPR array

For CRISPR-Cas types that depend on PAM sequences for target recognition, prespacer substrates at specific positions are processed to produce suitable PAMs. While each of the four Cas1 proteins contains a PAM-sensing domain, one is sufficient (Wang et al., 2015). Furthermore, the presence of a PAM in the prespacer can also ensure integration of the CRISPR array in the correct orientation (Shmakov et al., 2014). The prespacer processed by Cas1 then creates the two 3'-OH ends necessary for nucleophilic attack on the junction between the repeat and other sequences. The first nucleophilic attack most likely occurs at the leader-repeat junction and a second attack at the existing repeat-spacer junction (Jackson et al., 2017). In brief, Cas1-Cas2 depending on

IHF or LAS binds to the leader-proximal repeat, which helps junction between the leader and the repeat, or repeat and spacer, finally, the DNA repair mechanisms fill in the integration site (Figure 3).

4.2 crRNA biogenesis

During the second phase, crRNA biogenesis pathways diverge among the different CRISPR-Cas types. For example, in most type I and type III systems, Cas6 is essential for the primary processing of pre-crRNA (Makarova et al., 2015). Cas6 is a metal-independent endoribonuclease that recognizes and cleaves a single phosphodiester bond in the repeat sequences of pre-crRNA transcripts (Marraffini and Sontheimer, 2010). Type I-C systems do not contain Cas6 but instead contain the metal-independent endoribonuclease Cas5. For example, Cas5c in *S. pyogenes* not only processes pre-crRNA but also binds to the crRNA 5' handle, which is similar in function to Cas6 and Cas5 from type I-E systems in *E. coli*. Cas5d from the type I-C system in *S. mutans* UA159 also has RNase activity. Although present in types I-A, I-B and I-E, Cas5 is only directly involved in the maturation process in type I-C systems. However, type II-A systems *in S. pyogenes* use a distinct mechanism for crRNA biogenesis and this process depends on a tracrRNA and a host RNase III. The tracrRNA contains extensive secondary structure that is recognized and bound by Cas9 and also harbors a sequence that is complementary to the repeat sequence of the pre-crRNA. Annealing of these complementary sequences results in the formation of double stranded RNAs that are cleaved at one end by RNase III (Jinek et al., 2012). This cleavage liberates the small crRNAs from the precursor, which remain bound to Cas9 via their association with the tracrRNA (Pennisi, 2013).

4.3 Target interference

In the final phase, target interference, each crRNA combines with Cas protein(s) to form an effector complex, which is guided to target specific sequences (DNA or RNA) for recognition and destruction. In most cases, selection of an appropriate target sequence isn't entirely complementary to the crRNA, but is complementary to a so-called seed sequence (about seven or eight base pairs in close proximity to the PAM) (Sternberg et al., 2014). Several base mismatches between the spacer and protospacer are often tolerated(Semenova et al., 2011).

In type I systems, the Cascade complex scans DNA for complementary target sites, firstly identifying an appropriate PAM motif, then by partial melting and base pairing with the guide's seed sequence and eventually the formation of a complete R-loop structure, resulting in DNA degradation by the Cas3 nuclease/helicase (Jore et al., 2011). While type III systems are structurally similar to the type I systems (Jackson and Wiedenheft, 2015), there are some substantial mechanistic variations. Recently, it has been demonstrated that type III complexes are transcription-dependent DNA nucleases (Tamulaitis et al., 2014) meaning that they initially recognize target sequences via a specific interaction between the crRNA and a complementary nascent mRNA, following which cleavage of the flanking DNA sequences occurs. However, class 2 systems need only a single protein for interference. In the type II systems, the crRNA complex, which is involved in target recognition and cleavage, consists of Cas9 bound to a crRNA guide base paired with the tracrRNA. The crystal

structures of Cas9 show that the protein has two distinct roles in target recognition and nuclease activity. Finally, binding and cleavage of the target DNA by the Cas9-crRNA effector complex depends upon the recognition of an appropriate PAM located at the 3' end of the protospacer. Cas9 generates a blunt-ended double-strand break, typically located 3 nucleotides from the 3' end of the protospacer (Garneau et al., 2010).

5. Applications of CRISPR-Cas systems

5.1 Gene therapy

Genome editing using CRISPR-Cas9 technology is revolutionizing studies of gene functions and likely will give rise to an entirely new class of therapeutics for a wide range of diseases. One such study has successfully removed a mutation in the dystrophin protein, which is responsible for the most common form of Duchenne's muscular dystrophy (Dmd) (Nelson et al., 2016). Furthermore, CRISPR-Cas systems have been used in novel therapeutic approaches against human immunodeficiency virus type-1 (HIV-1), aiming to permanently inactivate all virus genomes or to prevent viral persistence in latent reservoirs (Wang et al., 2018). Although it remains to be seen which disease indications are most suited for CRISPR-Cas9-mediated genome editing, optimization of delivery and assessment of specificity are critical for the safe and effective translation of this technology to the clinical, especially in terms of off-target effects.

5.2 Transcriptional modulation

To repurpose the CRISPR-Cas system for genome regulation as opposed to genome editing, investigators have found that a catalytically inactive version of

Cas9 can be used as a platform for RNA-guided transcriptional regulation. A RNA-guided DNA recognition platform using co-expression of a mutant Cas9 with a small-guided (sgRNA) designed with a ~20 base-pair (bp) complementary region that yields specific silencing of a gene of interest without off-target effects has been developed and is known as CRISPR interference (CRISPRi) (Qi et al., 2013). This minimal CRISPRi system can not only effectively silence transcription initiation and elongation but is also highly specific and efficient with up to 1,000-fold repression. Moreover, the CRISPRi gene knockdown system is inducible and reversible and can be used to simultaneously regulate multiple genes, meaning that the CRISPRi system has the potential for genome-scale analysis and regulation of microbial genomes (Qi et al., 2012) and can be utilized as a flexible framework for engineering transcriptional regulatory networks. Finally, the CRISPRi system can also be used to knock down gene expression in mammalian cells (Qi et al., 2013).

5.3 Functional genomic screens

The high efficiency of gene editing based on the CRISPR-Cas system makes it possible to enable thorough genome-wide functional screens to identify essential or non-essential genes. Genome-wide loss-of-function screens using RNAi have many disadvantages, including partial knockdown of genes and serious off-target effects. However, Cas9-mediated sgRNA screens have high screening sensitivity and practicality and can be easily designed to target almost any DNA fragment. For example, high-throughput CRISPRi phenotyping has identified new essential genes in *S. pneumoniae* through the construction of a knockdown library targeting potentially essential genes by CRISPR interference and showing

a growth phenotype for most of them (Liu et al., 2017). Future application should be involve using multiplex sgRNA libraries to perturb noncoding genetic elements, distant enhancers, general promoter architectures or other transcriptional regulatory factors (Shalem et al., 2014).

5.4 Rapid cell generation and animal models

In the past few years, studies have shown that gene editing based on the CRISPR-Cas system can be used to accelerate the generation of transgenic models. To generate new cellular models, the sgRNA is designed from the sequence of interest, which is cloned into a plasmid encoding Cas9 and this plasmid is introduced into the target cells through transient transfection. For the generation of transgenic animal models, a previous study showed that Cas9 and the desired sgRNA can be directly injected into fertilized zygotes to obtain heritable gene modification in animals such as monkeys (Yang et al., 2013). Double-stranded DNA breaks (DSB) produced by Cas9 at specific genomic loci are repaired by the error-prone non-homologous end-joining (NHEJ) pathway generating insertions or deletions, or by homology-directed repair (HDR) using a homologous DNA repair template to generate the desired mutants (Salsman and Dellaire, 2017). Furthermore, CRISPRi can also be used to repress gene expression to identify its functions, rather than produce DSB. Therefore, based on these strategies, new transgenic animal models can be constructed. Comparison with the traditional process of generating transgenic animals, the generation time for mutant models can be reduced from more than a year to only several weeks, saving costs and enlarging the scale at which these experiments can be performed. More interestingly, Cas9 can also be used for direct

modification of somatic tissue to bypass embryonic manipulation (Hsu et al., 2014).

5.5 Using CRISPR-Cas systems as antimicrobials

Regulating the composition of microbial populations plays a critical role in medicine and biotechnology. While traditional strategies such as antibiotics and antimicrobial peptides provide partial solutions, these have many disadvantages, such as lack of specificity or high cost. A study demonstrated that RNA-directed CRISPR-Cas systems as antimicrobials can overcome those drawbacks (Gomaa et al., 2014). Although the CRISPR-Cas system naturally evolved to provide adaptive immunity against invasive genetic elements in bacteria and archaea, genome editing technologies using this system can be co-opted to target chromosomal sequences, also referred as 'self-targeting'. However, self-targeting by CRISPR-Cas systems is typically lethal in bacteria if DNA repair is absent (Bikard and Barrangou, 2017). For example, studies have found that crRNAs guide Cas9 to cleave all copies of the genome in *E. coli* so efficiently that it is impossible to make repair these breaks through recombination with a sister chromosome, resulting in cell death (Cui and Bikard, 2016). However, a major limit remains in the effective and steady delivery of a CRISPR-Cas9 system into host cells, especially in Gram-positive bacteria, although several strategies using phages, lentivirus (LV), adenovirus (AV) and adeno-associated virus (AAV) as vectors have been successful (Wang et al., 2017). Future research should focus on CRISPR-Cas and delivery system combinations to develop specific and programmable antimicrobials that will open new avenues for the application of CRISPR-Cas9 systems to effectively kill pathogens and precisely alter

microbiome composition (Bikard and Barrangou, 2017; Citorik et al., 2014).

5.6 Anti-CRISPR

Like the human immune system, the CRISPR-Cas system can rapidly recognize and clear invading genetic elements in bacteria. However, some phages have also developed a way to escape recognition by CRISPR-Cas systems. These evasion strategies contain point mutations, large-scale deletions, DNA modifications or phage-encoded proteins known as anti-CRISPRs (Acrs) (Hynes et al., 2017). To date, 21 unique families of anti-CRISPR proteins against type I and type II CRISPR–Cas systems have been discovered. For example, AcrF1 and AcrF2 were shown to interact directly with type I-F Cascade complex in *P. aeruginosa*, inhibiting its ability to bind with DNA (Bondy-Denomy et al., 2013). Moreover, AcrIIA4 abolishes the type II-A CRISPR-Cas9 system in *S. pyogenes* by modeling DNA sequence and occupying the PAM site bound by Cas9, resulting in the failure of Cas9 to bind target DNA (Dong et al., 2017). However, hosts can also acquire more new spacers derived from viruses in response to these escape mutations (Fineran et al., 2014). Therefore, there is an evolutionary arms race between anti-CRISPR and CRISPR-Cas systems. The presence of anti-CRISPRs also contributes to CRISPR-Cas system diversity, which provides strong selective pressure on the evolution of anti-CRISPR proteins. Two anti-CRISPRs derived from *L. monocytogenes* have the ability to prevent genome editing commonly used by *S. pyogenes* Cas9 (SpyCas9) in human cells (Rauch et al., 2017), which may act as an ‘off-switch’ to control the activity of CRISPR-Cas systems. Moreover, as there are many human pathogens with CRISPR-Cas systems, phages used for gene therapy could be

equipped with a variety of anti-CRISPRs to prevent the bacterial adaptive immune response. Finally, our understanding of anti-CRISPRs in its infancy, but these studies have opened up a new area of CRISPR-Cas research and subsequent studies will uncover more mechanisms to enrich CRISPR-Cas system diversity (Maxwell, 2017).

5.7 Other applications

To study real-time interactions between specific genes, robust strategies are required to visualize DNA in living cells. Traditional methods (such as fluorescence *in situ* hybridization, FISH) for labeling DNA require sample fixation and can't capture all live processes. Cas9 tagged with fluorescent labels for labeling specific DNA loci was recently developed as a powerful tool for live-cell-imaging and is known as DNA-FISH (Chen et al., 2013). Moreover, the application of CRISPR-Cas systems for molecular detection includes the detection of specific DNA or RNA sequences, detection of mutations and single nucleotide variants (SNVs), and the detection of pathogens (Zhou et al., 2018). To date, the majority of CRISPR-based applications have focused on the type II-A CRISPR-Cas9 system in *S. pyogenes*, whereas a significant diversity in CRISPR types remains to be utilized, expanding their potential applications.

6. Conclusion

This paper has reviewed current advances in our understanding of CRISPR-Cas systems in streptococci, and has looked at the potential for the application of this technology to other fields. Interestingly, those microbes that are responsible for some of the most prevalent infectious diseases in the world may also harbor a

cure for a number of genetic diseases and may provide new therapies to combat diseases and new genetic tools that can revolutionize medicine. However, the applications of CRISPR-Cas systems face several challenges, including off-target effects, delivery efficiency and unwanted mutations. Thus, future work is required to discovering other novel CRISPR-Cas systems, overcome current limitations and address the functions of the CRISPR-Cas locus in ecology, evolution and pathogenesis.

7. Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (grant numbers 31200985, 81430011, and 31870065). The authors declare no potential conflicts of interest with regard to the authorship and/or publication of this article.

8. References

Abu Kwaik Y, Gao LY, Stone BJ, Venkataraman C, Harb OS. (1998). Invasion of protozoa by Legionella pneumophila and its role in bacterial ecology and pathogenesis. Appl Environ Microbiol. 64(9):3127-3133. PMID: 9726849

Amitai G, Sorek R. (2016). CRISPR-Cas adaptation: insights into the mechanism of action. Nat Rev Microbiol. 14(2):67-76. DOI: 10.1038/nrmicro.2015.14.

Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, Gagarinova A, Pogoutse O, Brown G, Binkowski A, et al. (2011). A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Mol Microbiol. 79(2):484-502. DOI: 10.1111/j.1365-2958.2010.07465.x.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S,

Romero DA, Horvath P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science. 315(5819):1709-1712. DOI: 10.1126/science.1138140.

Bethke J, Avendaño-Herrera R. (2017). Comparative genome analysis of two *Streptococcus phocae* subspecies provides novel insights into pathogenicity. Mar Genomics. 31:53-61. DOI: 10.1016/j.margen.2016.09.005.

Bikard D, Barrangou R. (2017). Using CRISPR-Cas systems as antimicrobials. Curr Opin Microbiol. 37:155-160. DOI: 10.1016/j.mib.2017.08.005.

Biswas A, Gagnon JN, Brouns SJ, Fineran PC, Brown CM. (2013). CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. RNA Biol. 10(5):817-827. DOI: 10.4161/rna.24046.

Biswas A, Staals RH, Morales SE, Fineran PC, Brown CM. (2016). CRISPRDetect: A flexible algorithm to define CRISPR arrays. BMC Genomics. 17:356. DOI: 10.1186/s12864-016-2627-0.

Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 493(7432):429-432. DOI: 10.1038/nature11723.

Cady KC, O'Toole GA. (2011). Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. J Bacteriol. 193(14):3433-3445. DOI: 10.1128/JB.01411-10.

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 155(7):1479-1491. DOI: 10.1016/j.cell.2013.12.001.

Chen J, Li T, Zhou X, Cheng L, Huo Y, Zou J, Li Y. (2017). Characterization of the

clustered regularly interspaced short palindromic repeats sites in Streptococcus mutans isolated from early childhood caries patients. Arch Oral Biol. 83:174-180. DOI: 10.1016/j.archoralbio.2017.07.023.

Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, Siepel A. (2012). Replacing and additive horizontal gene transfer in Streptococcus. Mol Biol Evol. 29(11):3309-3320. DOI: 10.1093/molbev/mss138.

Chylinski K, Le Rhun A, Charpentier E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol. 10(5):726-737. DOI: 10.4161/rna.24321.

Chylinski K, Makarova KS, Charpentier E, Koonin EV. (2014). Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Res. 42(10):6091-6105. DOI: 10.1093/nar/gku241.

Citorik RJ, Mimee M, Lu TK. (2014). Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol. 32(11):1141-1145. DOI: 10.1038/nbt.3011.

Cui L, Bikard D. (2016). Consequences of Cas9 cleavage in the chromosome of Escherichia coli. Nucleic Acids Res. 44(9):4243-51. DOI: 10.1093/nar/gkw223.

Dong, Guo M, Wang S, Zhu Y, Wang S, Xiong Z, Yang J, Xu Z, Huang Z. (2017). Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. Nature. 546(7658):436-439. DOI: 10.1038/nature22377.

Facklam R. (2002). What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin Microbiol Rev. 15(4):613-630. DOI: 10.1128/CMR.15.4.613-630.2002.

Fineran PC, Gerritzen MJ, Suárez-Diez M, Künne T, Boekhorst J, van Hijum SA, Staals RH, Brouns SJ. (2014). Degenerate target sites mediate rapid primed

CRISPR adaptation. Proc Natl Acad Sci U S A. 111(16):E1629-38. DOI: 10.1073/pnas.1400071111.

Fischetti VA. (2007). In vivo acquisition of prophage in Streptococcus pyogenes. Trends Microbiol. 15(7):297-300. DOI: 10.1016/j.tim.2007.05.003.

Fontaine L, Wahl A, Fléchard M, Mignolet J, Hols P. (2015). Regulation of competence for natural transformation in streptococci. Infect Genet Evol. 33:343-360. DOI: 10.1016/j.meegid.2014.09.010.

Garneau JE, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 468(7320):67-71. DOI: 10.1038/nature09523.

Gasiunas G, Barrangou R, Horvath P, Siksnys V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A. 109(39):E2579-86. DOI: 10.1073/pnas.1208507109.

Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. (2014). Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. MBio. 5(1):e00928-13. DOI: 10.1128/mBio.00928-13.

Grissa I, Vergnaud G, Pourcel C. (2007a). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics. 8:172. DOI: 10.1186/1471-2105-8-172.

Grissa I, Vergnaud G, Pourcel C. (2007b). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35(Web Server issue):W52-7. DOI: 10.1093/nar/gkm360.

Gunderson FF, Cianciotto NP. (2013). The CRISPR-associated gene cas2 of

Legionella pneumophila is required for intracellular infection of amoebae. MBio. 4(2):e00074-13. DOI: 10.1128/mBio.00074-13.

Hamada S, Kawabata S, Nakagawa I. (2015). Molecular and genomic characterization of pathogenic traits of group A *Streptococcus pyogenes*. Proc Jpn Acad Ser B Phys Biol Sci. 91(10):539-559. DOI: 10.2183/pjab.91.539.

Heler R, Marraffini LA, Bikard D. (2014). Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. Mol Microbiol. 93(1):1-9. DOI: 10.1111/mmi.12640.

Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, Marraffini LA. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature. 519(7542):199-202. DOI: 10.1038/nature14245.

Hols P, Hancy F, Fontaine L, Grossiord B, Prozzi D, Leblond-Bourget N, Decaris B, Bolotin A, Delorme C, Dusko Ehrlich S, et al. (2005). New insights in the molecular biology and physiology of Streptococcus thermophilus revealed by comparative genomics. FEMS Microbiol Rev. 29(3):435-463. DOI: 10.1016/j.fmrre.2005.04.008.

Hossain M, Egan SA, Coffey T, Ward PN, Wilson R, Leigh JA, Emes RD. (2015). Virulence related sequences; insights provided by comparative genomics of *Streptococcus uberis* of differing virulence. BMC Genomics. 16:334. DOI: 10.1186/s12864-015-1512-6.

Hsu PD, Lander ES, Zhang F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell. 157(6):1262-1278. DOI: 10.1016/j.cell.2014.05.010.

Huo Y, Nam KH, Ding F, Lee H, Wu L, Xiao Y, Farchione MD Jr, Zhou S, Rajashankar K, Kurinov I, et al. (2014). Structures of CRISPR Cas3 offer

mechanistic insights into Cascade-activated DNA unwinding and degradation. Nat Struct Mol Biol. 21(9):771-777. DOI: 10.1038/nsmb.2875.

Hynes AP, Rousseau GM, Lemay ML, Horvath P, Romero DA, Fremaux C, Moineau S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. Nat Microbiol. 2(10):1374-1380. DOI: 10.1038/s41564-017-0004-7.

Ivančić-Baće I, Cass SD, Wearne SJ, Bolt EL. (2015). Different genome stability proteins underpin primed and naïve adaptation in *E. coli* CRISPR-Cas immunity. Nucleic Acids Res. 43(22):10821-10830. DOI: 10.1093/nar/gkv1213.

Jackson RN, Wiedenheft B. (2015). A conserved structural chassis for mounting versatile CRISPR RNA-Guided immune responses. Mol Cell. 58(5):722-728. DOI: 10.1016/j.molcel.2015.05.023.

Jackson SA, McKenzie RE, Fagerlund RD, Kieper SN, Fineran PC, Brouns SJ. (2017). CRISPR-Cas: Adapting to change. Science. 356(6333). pii: eaal5056. DOI: 10.1126/science.aal5056.

Jans C, Follador R, Hochstrasser M, Lacroix C, Meile L, Stevens MJ. (2013). Comparative genome analysis of Streptococcus infantarius subsp. infantarius CJ18, an African fermented camel milk isolate with adaptations to dairy environment. BMC Genomics. 14:200. DOI: 10.1186/1471-2164-14-200.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 337(6096):816-821. DOI: 10.1126/science.1225829.

Jones CL, Sampson TR, Nakaya HI, Pulendran B, Weiss DS. (2012). Repression of bacterial lipoprotein production by Francisella novicida facilitates evasion of innate immune recognition. Cell Microbiol. 14(10):1531-1543. DOI:

10.1111/j.1462-5822.2012.01816.x.

Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul U, Wurm R, Wagner R, et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat Struct Mol Biol. 18(5):529-536. DOI: 10.1038/nsmb.2019.

Jung CJ, Yeh CY, Shun CT, Hsu RB, Cheng HW, Lin CS, Chia JS. (2012). Platelets enhance biofilm formation and resistance of endocarditis-inducing streptococci on the injured heart valve. J Infect Dis. 205(7):1066-1075. DOI: 10.1093/infdis/jis021.

Kilian M, Riley DR, Jensen A, Brüggemann H, Tettelin H. (2014). Parallel evolution of Streptococcus pneumoniae and Streptococcus mitis to pathogenic and mutualistic lifestyles. MBio. 5(4):e01490-14. DOI: 10.1128/mBio.01490-14.

Koonin EV, Makarova KS, Zhang F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. Curr Opin Microbiol. 37:67-78. DOI: 10.1016/j.mib.2017.05.008.

Leenay RT, Beisel CL. (2017). Deciphering, Communicating, and Engineering the CRISPR PAM. J Mol Biol. 2017 Jan 20;429(2):177-191. DOI: 10.1016/j.jmb.2016.11.024.

Lier C, Baticle E, Horvath P, Haguenoer E, Valentin AS, Glaser P, Mereghetti L, Lanotte P. (2015). Analysis of the type II-A CRISPR-Cas system of *Streptococcus agalactiae* reveals distinctive features according to genetic lineages. Front Genet. 6:214. DOI: 10.3389/fgene.2015.00214.

Liu X, Gallay C, Kjos M, Domenech A, Slager J, van Kessel SP, Knoops K, Sorg RA, Zhang JR, Veening JW. (2017). High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. Mol Syst Biol.

13(5):931. DOI: 10.15252/msb.20167449.

Lopez-Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina E, Gonzalez-Zorn B, Poyart C, Rosinski-Chupin I, Glaser P. (2012). The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. Mol Microbiol. 85(6):1057-1071. DOI: 10.1111/j.1365-2958.2012.08172.x.

Louthrenoo W, Kasitanon N, Wangkaew S, Hongsongkiat S, Sukitawut W, Wichainun R. (2014). Streptococcus agalactiae: an emerging cause of septic arthritis. J Clin Rheumatol. 20(2):74-78. DOI: 10.1097/RHU.0000000000000071.

Louwen R, Staals RH, Endtz HP, van Baarlen P, van der Oost J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. Microbiol Mol Biol Rev. 78(1):74-88. DOI: 10.1128/MMBR.00039-13.

Lynch JP, Zhanel GG. (2009). Streptococcus pneumoniae: epidemiology, risk factors, and strategies for prevention. Semin Respir Crit Care Med. 30(2):189-209. DOI: 10.1055/s-0029-1202938.

Maeda Y, Goldsmith CE, Coulter WA, Mason C, Dooley JS, Lowery CJ, Snelling WJ, Moore JE. (2011). Prevalence of clustered regulatory interspaced short palindromic repeat (CRISPR)-like sequences in mitis-group streptococci. Br J Biomed Sci. 68(2):65-8. PMID: 21706916

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol. 9(6):467-477. DOI: 10.1038/nrmicro2577.

Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH, et al. (2015). An updated

evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol. 13(11):722-736. DOI: 10.1038/nrmicro3569.

Makarova KS, Wolf YI, Koonin EV. (2013). The basic building blocks and evolution of CRISPR-CAS systems. Biochem Soc Trans. 41(6):1392-1400. DOI: 10.1042/BST20130038.

Marraffini LA, Sontheimer EJ. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet. 11(3):181-190. DOI: 10.1038/nrg2749.

Marraffini LA. (2015). CRISPR-Cas immunity in prokaryotes. Nature. 526(7571):55-61. DOI: 10.1038/nature15386.

Maxwell KL. (2017). The Anti-CRISPR Story: A Battle for Survival. Mol Cell. 68(1):8-14. DOI: 10.1016/j.molcel.2017.09.002.

Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Castellanos Rivera RM, Madhavan S, Pan X, Ran FA, Yan WX, et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 351(6271):403-407. DOI: 10.1126/science.aad5143.

Nho SW, Hikima J, Park SB, Jang HB, Cha IS, Yasuike M, Nakamura Y, Fujiwara A, Sano M, Kanai K, et al. (2013). Comparative genomic characterization of three *Streptococcus parauberis* strains in fish pathogen, as assessed by wide-genome analyses. PLoS One. 8(11):e80395. DOI: 10.1371/journal.pone.0080395.

Nuñez JK, Bai L, Harrington LB, Hinder TL, Doudna JA. (2016). CRISPR immunological memory requires a host factor for specificity. Mol Cell. 62(6):824-833. DOI: 10.1016/j.molcel.2016.04.027.

Nuñez JK, Harrington LB, Kranzusch PJ, Engelman AN, Doudna JA. (2015). Foreign DNA capture during CRISPR-Cas adaptive immunity. Nature.

527(7579):535-538. DOI: 10.1038/nature15760.

Okamoto M, Naito M, Miyanohara M, Imai S, Nomura Y, Saito W, Momoi Y, Takada K, Miyabe-Nishiwaki T, Tomonaga M, et al. (2016). Complete genome sequence of Streptococcus troglodytae TKU31 isolated from the oral cavity of a chimpanzee (Pan troglodytes). Microbiol Immunol. 60(12):811-816. DOI: 10.1111/1348-0421.12453.

Olson AB, Kent H, Sibley CD, Grinwis ME, Mabon P, Ouellette C, Tyson S, Graham M, Tyler SD, Van Domselaar G, et al. (2013). Phylogenetic relationship and virulence inference of *Streptococcus anginosus* Group: curated annotation and whole-genome comparative analysis support distinct species designation. BMC Genomics. 14:895. DOI: 10.1186/1471-2164-14-895.

Pennisi E. (2013). The CRISPR craze. Science. 341(6148):833-836. DOI: 10.1126/science.341.6148.833.

Qi L, Haurwitz RE, Shao W, Doudna JA, Arkin AP. (2012). RNA processing enables predictable programming of gene expression. Nat Biotechnol. 30(10):1002-1006. DOI: 10.1038/nbt.2355.

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 152(5):1173-1183. DOI: 10.1016/j.cell.2013.02.022.

Rauch BJ, Silvis MR, Hultquist JF, Waters CS, McGregor MJ, Krogan NJ, Bondy-Denomy J. (2017). Inhibition of CRISPR-Cas9 with bacteriophage proteins. Cell. 168(1-2):150-158.e10. DOI: 10.1016/j.cell.2016.12.009.

Salsman J, Dellaire G. (2017). Precision genome editing in the CRISPR era. Biochem Cell Biol. 95(2):187-201. DOI: 10.1139/bcb-2016-0137.

Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A, Marraffini LA. (2015). Co-transcriptional DNA and RNA cleavage during Type III CRISPR-Cas immunity. Cell. 161(5):1164-1174. DOI: 10.1016/j.cell.2015.04.027.

Sampson TR, Weiss DS. (2014). CRISPR-Cas systems: new players in gene regulation and bacterial physiology. Front Cell Infect Microbiol. 4:37. DOI: 10.3389/fcimb.2014.00037.

Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJ, 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(25):10098-10103. DOI: 10.1073/pnas.1104144108.

Serbanescu MA, Cordova M, Krastel K, Flick R, Beloglazova N, Latos A, Yakunin AF, Senadheera DB, Cvitkovitch DG. Role of the Streptococcus mutans CRISPR-Cas systems in immunity and cell physiology. J Bacteriol. 197(4):749-761. DOI: 10.1128/JB.02333-14.

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 343(6166):84-87. DOI: 10.1126/science.1247005.

Shimomura Y, Okumura K, Murayama SY, Yagi J, Ubukata K, Kirikae T, Miyoshi-Akiyama T. (2011). Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. equisimilis strain causing streptococcal toxic shock syndrome (STSS). BMC Genomics. 12:17. DOI: 10.1186/1471-2164-12-17.

Shmakov S, Savitskaya E, Semenova E, Logacheva MD, Datsenko KA,

Severinov K. (2014). Pervasive generation of oppositely oriented spacers during CRISPR adaptation. Nucleic Acids Res. 42(9):5907-5916. DOI: 10.1093/nar/gku226.

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 507(7490):62-67. DOI: 10.1038/nature13011.

Tamulaitis G, Kazlauskiene M, Manakova E, Venclovas Č, Nwokeoji AO, Dickman MJ, Horvath P, Siksnys V. (2014). Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. Mol Cell. 56(4):506-517. DOI: 10.1016/j.molcel.2014.09.027.

Tamulaitis G, Venclovas Č, Siksnys V. (2017). Type III CRISPR-Cas immunity: major differences brushed aside. Trends Microbiol. 25(1):49-61. DOI: 10.1016/j.tim.2016.09.012.

Tanzer JM, Livingston J, Thompson AM. (2001). The microbiology of primary dental caries in humans. J Dent Educ. 65(10):1028-1037. PMID: 11699974

Tohya M, Sekizaki T, Miyoshi-Akiyama T. (2018). Complete genome sequence of *Streptococcus ruminantium* sp. nov. GUT-187T (=DSM 104980T =JCM 31869T), the type strain of *S. ruminantium*, and comparison with genome sequences of *Streptococcus suis* strains. Genome Biol Evol. 10(4):1180-1184. DOI: 10.1093/gbe/evy078.

Vestergaard G, Garrett RA, Shah SA. (2014). CRISPR adaptive immune systems of Archaea. RNA Biol. 11(2):156-167. DOI: 10.4161/rna.27990.

Waller AS, Robinson C. (2013). *Streptococcus zooepidemicus* and *Streptococcus equi* evolution: the role of CRISPRs. Biochem Soc Trans. 2013 Dec;41(6):1437-1443. DOI: 10.1042/BST20130165.

Wang G, Zhao N, Berkhout B, Das AT. (2018). CRISPR-Cas based antiviral strategies against HIV-1. Virus Res. 244:321-332. DOI: 10.1016/j.virusres.2017.07.020.

Wang HX, Li M, Lee CM, Chakraborty S, Kim HW, Bao G, Leong KW. (2017). CRISPR/Cas9-Based genome editing for disease modeling and therapy: challenges and opportunities for nonviral delivery. Chem Rev. 117(15):9874-9906. DOI: 10.1021/acs.chemrev.6b00799.

Wang J, Li J, Zhao H, Sheng G, Wang M, Yin M, Wang Y. (2015). Structural and sechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. Cell. 163(4):840-853. DOI: 10.1016/j.cell.2015.10.008.

Wei Y, Chesne MT, Terns RM, Terns MP. (2015). Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in Streptococcus thermophilus. Nucleic Acids Res. 43(3):1749-1758. DOI: 10.1093/nar/gku1407

Wei Y, Chesne MT, Terns RM, Terns MP. (2015). Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus therm*ophilus. Nucleic Acids Res. 43(3):1749-1758. DOI:10.1093/nar/gku1407.

Willenborg J, Goethe R. (2016). Metabolic traits of pathogenic streptococci. FEBS Lett. 590(21):3905-3919. DOI: 10.1002/1873-3468.12317.

Wright AV, Doudna JA. (2016). Protecting genome integrity during CRISPR immune adaptation. Nat Struct Mol Biol. 23(10):876-883. DOI: 10.1038/nsmb.3289.

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell. 154(6):1370-1379. DOI: 10.1016/j.cell.2013.08.022.

Yao X, Li M, Wang J, Wang C, Hu D, Zheng F, Pan X, Tan Y, Zhao Y, Hu L, Tang J, Hu F. (2015). Isolation and characterization of a native avirulent strain of *Streptococcus suis* serotype 2: a perspective for vaccine development. Sci Rep. 5:9835. DOI: 10.1038/srep09835.

Yombi Jc, Belkhir L, Jonckheere S, Wilmes D, Cornu O, Vandercam B, Rodriguez-Villalobos H. (2012). Streptococcus gordonii septic arthritis: two cases and review of literature. BMC Infect Dis. 12:215. DOI: 10.1186/1471-2334-12-215.

Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of Pseudomonas aeruginosa. J Bacteriol. 191(1):210-219. DOI: 10.1128/JB.00797-08.

Zhang BC, Zhang J, Sun L. (2014). *Streptococcus iniae* SF1: complete genome sequence, proteomic profile, and immunoprotective antigens. PLoS One. 9(3):e91324. DOI: 10.1371/journal.pone.0091324.

Zhou L, Peng R, Zhang R, Li J. (2018). The applications of CRISPR/Cas system in molecular detection. J Cell Mol Med. 22(12):5807-5815. DOI: 10.1111/jcmm.13925.