



Endogenous Type I CRISPR-Cas: From Foreign DNA Defense to Prokaryotic Engineering

Yanli Zheng, Jie Li, Baiyang Wang, Jiamei Han, Yile Hao, Shengchen Wang, Xiangdong Ma, Shihui Yang, Lixin Ma, Li Yi* and Wenfang Peng*

State Key Laboratory of Biocatalysis and Enzyme Engineering, Hubei Collaborative Innovation Center for Green Transformation of Bio-resources, Hubei Key Laboratory of Industrial Biotechnology, School of Life Sciences, Hubei University, Wuhan, China

OPEN ACCESS

Edited by:

C. Perry Chou,
University of Waterloo, Canada

Reviewed by:

Scott Bailey,
Johns Hopkins University,
United States
Matthew Wook Chang,
National University of Singapore,
Singapore

*Correspondence:

Li Yi
liy@hubu.edu.cn
Wenfang Peng
wenfang@hubu.edu.cn

Specialty section:

This article was submitted to
Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 30 October 2019

Accepted: 24 January 2020

Published: 04 March 2020

Citation:

Zheng Y, Li J, Wang B, Han J,
Hao Y, Wang S, Ma X, Yang S, Ma L,
Yi L and Peng W (2020) Endogenous
Type I CRISPR-Cas: From Foreign
DNA Defense to Prokaryotic
Engineering.
Front. Bioeng. Biotechnol. 8:62.
doi: 10.3389/fbioe.2020.00062

Establishment of production platforms through prokaryotic engineering in microbial organisms would be one of the most efficient means for chemicals, protein, and biofuels production. Despite the fact that CRISPR (clustered regularly interspaced short palindromic repeats)-based technologies have readily emerged as powerful and versatile tools for genetic manipulations, their applications are generally limited in prokaryotes, possibly owing to the large size and severe cytotoxicity of the heterogeneous Cas (CRISPR-associated) effector. Nevertheless, the rich natural occurrence of CRISPR-Cas systems in many bacteria and most archaea holds great potential for endogenous CRISPR-based prokaryotic engineering. The endogenous CRISPR-Cas systems, with type I systems that constitute the most abundant and diverse group, would be repurposed as genetic manipulation tools once they are identified and characterized as functional in their native hosts. This article reviews the major progress made in understanding the mechanisms of invading DNA immunity by type I CRISPR-Cas and summarizes the practical applications of endogenous type I CRISPR-based toolkits for prokaryotic engineering.

Keywords: endogenous type I CRISPR-Cas systems, mechanisms of action, DNA targeting, genome editing, selective killing, antimicrobials, gene expression modulation

INTRODUCTION

Throughout the past billion years, bacteria and archaea have evolved a range of defense mechanisms to defend themselves against their viral predators (Doron et al., 2018), including restriction-modification systems, abortive infections and phage adsorption blocks, and the recently discovered CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR-associated) systems (Jansen et al., 2002). Unique among these mechanisms, CRISPR immunity functions by storing records of previous invasions to provide immunological memory for a rapid and robust response upon subsequent viral infections.

CRISPR-Cas systems consist of two genetic components, the CRISPR array and *cas* genes encoding Cas proteins. The CRISPR array, featuring the CRISPR-Cas systems, is composed of conserved direct repeats, which are separated by unique sequences derived from the invasive mobile genetic elements (termed as *spacers*) (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). Generally, the CRISPR immunity is driven by the Cas proteins in three distinct molecular stages.

The first stage is termed *spacer acquisition*, in which a short DNA stretch is captured from an invading genetic element and incorporated into a CRISPR array as the first spacer immediately after a leader sequence. Then, in the process of *crRNA (CRISPR RNA) biogenesis*, the entire CRISPR array is transcribed into a precursor CRISPR RNA (pre-crRNA) molecule driven by promoter elements embedded in the leader sequence. Following the transcription, cleavage within the repeats of the pre-crRNA by ribonucleases gives mature crRNAs, with each carrying a unique foreign sequence. In the final stage of *target interference*, each crRNA forms a ribonucleoprotein effector complex with Cas proteins and guides the effector machinery to the matching region of the invader through base pairing for destruction (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008; Garneau et al., 2010).

Despite the general immunity stages for all the characterized CRISPR-Cas systems, the Cas proteins and hence the effector complexes vary widely. According to the assortment of *cas* genes and the complexity of their inference complexes, to date, six main types, types I to VI, of CRISPR-Cas systems have been identified and grouped into two classes, classes 1 and 2 (Makarova et al., 2015, 2018; **Table 1**). Systems of class 1, including types I, III, and IV, are defined by multi-Cas proteins, whereas those of class 2, including types II, V, and VI, utilize a single effector Cas protein (**Figure 1A**).

Because of the simplicity of class 2 systems, in which a single Cas protein is sufficient to execute targets binding and cleavage, their effector machineries are relatively easy to be adopted and have been emerged as powerful tools for genome manipulation applications in both prokaryotic and eukaryotic cells. Type II CRISPR-Cas9, as the first identified class 2 system, has been extensively harnessed for genome editing in a wide range of organism, from bacteria to eukaryotic cells, in the past few years (Jinek et al., 2012; Cong et al., 2013; Malina et al., 2013; Jiang et al., 2014). Subsequently, type V CRISPR-Cas12 was characterized and repurposed for genome editing (Zetsche et al., 2015, 2017; Fonfara et al., 2016). These systems have been also engineered for applications beyond genome editing, for example, gene expression regulation via repression or activation, epigenome editing, *in situ* genomic imaging, large-scale genomic screening, and so on (Chen et al., 2013; Cheng et al., 2013; Perez-Pinera et al., 2013; Qi et al., 2013; Shalem et al., 2014; Takei et al., 2017). Type VI CRISPR-Cas13 is the third type in class 2 system, which was demonstrated and developed as a versatile RNA manipulation tool to be used in RNA interference (RNAi), *in vivo* RNA molecule visualization, and nucleic acid detection (Cox et al., 2017; O'Connell, 2019). Recently, several novel single-stranded DNA (ssDNA)-cleaving CRISPR-Cas14 systems belonging to class 2 were identified (Harrington et al., 2018). It is worth to be mentioned that the targeted non-specific ssDNA cleavage activity of Cas14 enabled the system to perform genotyping (Harrington et al., 2018), while it has only about half size of Cas9/Cpf1, thus representing so far the smallest effector nuclease in a single-Cas effector system.

Undoubtedly, class 2 CRISPR-Cas systems have attracted great attention with their fruitful achievements in genome manipulations. However, the engineered targets exhibited an

obvious bias toward eukaryotic cells. One of the reasons could be that, as heterologous large-sized nucleases with intrinsic toxicity, the class 2 Cas effectors are hard to be introduced into prokaryotes, particularly those poorly genetically accessible prokaryotic cells. In one recent study, it failed to yield any colony when introducing Cas9 into *Corynebacterium glutamicum* cells even without a guide RNA (gRNA), suggesting the cytotoxicity of the Cas9 *per se* (Jiang et al., 2017). It was also reported that overexpressing a catalytically dead Cas9 (dCas9) in *Escherichia coli* resulted in abnormal morphology and retarded growth, indicating that the cytotoxicity of Cas9 is not solely caused by DNA cleavage but possibly transient non-specific DNA binding across the genome (Cho et al., 2018). Therefore, using endogenous CRISPR-Cas systems of the host for genome engineering could be an effective way to overcome the restriction (Hidalgo-Cantabrana et al., 2019a). In comparison with the imported class 2 systems, all the protein components of endogenous type I systems are present in the cells, excluding any heterologous nuclease. They naturally produce mature crRNA guides without any heterologous helper, which is with particular convenience to conduct, for example, multiplexed genome editing or simultaneous multiple-gene regulation for metabolic pathway engineering by simply supplying an artificial CRISPR array (Luo et al., 2015; Cheng et al., 2017).

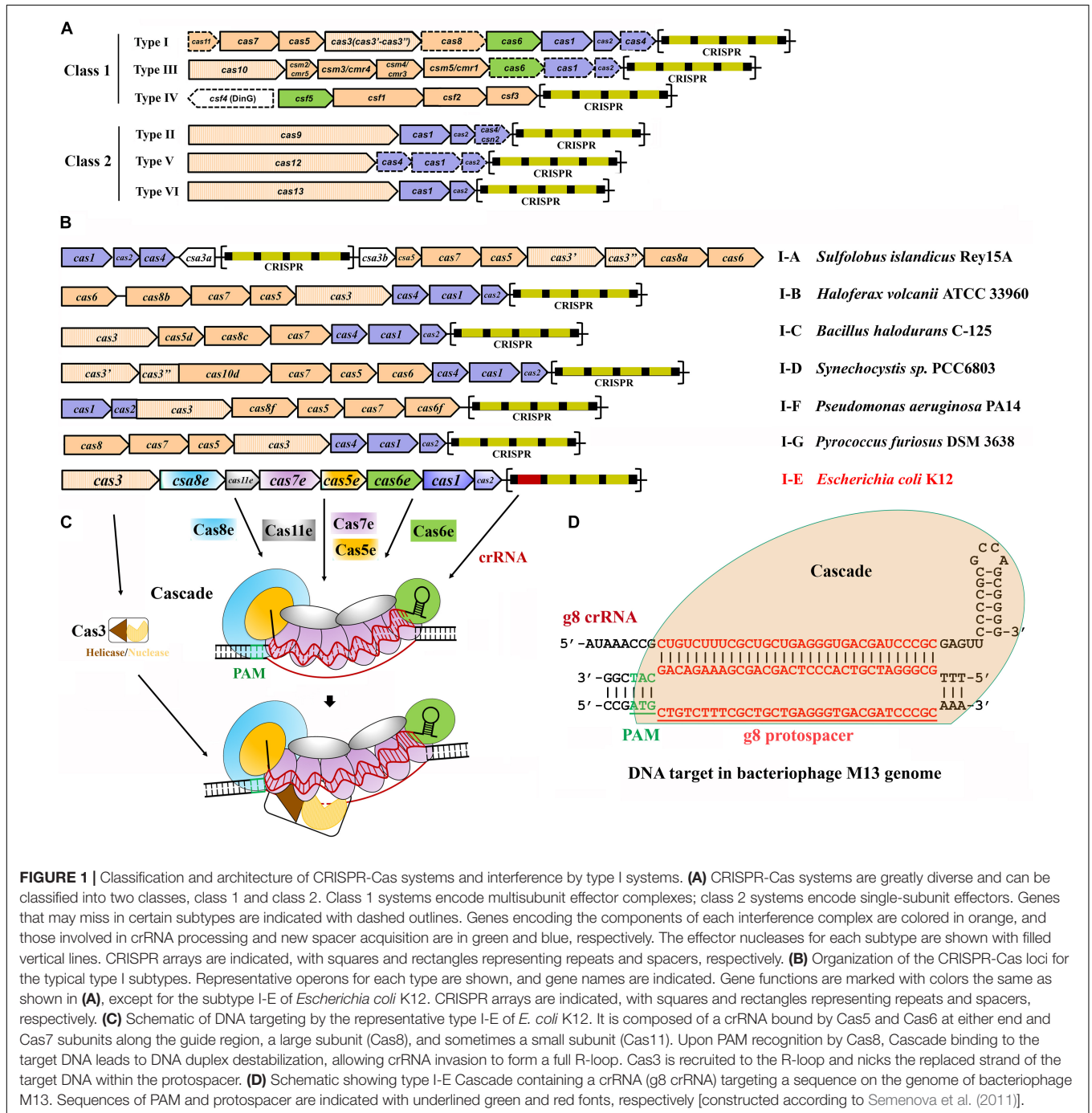
It was reported that class 1 systems, primarily types I and III, with only a few reports added recently of type IV (Crowley et al., 2019; Ozcan et al., 2019; Taylor et al., 2019), are present in more than 90% of sequenced genomes of bacteria and archaea (Makarova et al., 2015). Class 1 endogenous systems are much more abundant than class 2 systems (Grissa et al., 2007) and exist not only in mesophiles but also in extreme thermophiles. This is indicative of a great potential to harness the class 1 endogenous systems for applications across many areas of biology. In fact, using type III system for genome editing has been carried out in *Sulfolobus islandicus* (Li et al., 2016). Because the functions and applications of type III systems have been recently reviewed elsewhere (Liu et al., 2018), we will present here the major progress achieved for type I systems, uncovering mechanisms of action of the type I CRISPR-directed immunity, concerning crRNA processing, effector complex assembly, PAM-directed target recognition, seed sequence-mediated target invasion, and Cas3-executed target cleavage, and briefly discuss their recent practical applications, for example, genome editing, antimicrobials, and gene expression regulation, in their prokaryotic hosts.

MECHANISMS OF INVADING DNA DEFENSE BY TYPE I CRISPR-CAS

Type I CRISPR-Cas systems contain a Cas3 protein as the defining feature of this type with exception for transposon-encoded type I variants (Peters et al., 2017; Klompe et al., 2019; Makarova et al., 2019; Strecker et al., 2019) and are divided into six subtypes, I-A through I-G, with each harboring a specifying Cas8 homolog (**Figure 1B** and **Table 1**; Majumdar et al., 2015; Makarova et al., 2015, 2018; Koonin et al., 2017b).

TABLE 1 | The six types of CRISPR-Cas systems.

Class	Type	Subtype	Spacer acquisition	crRNA biogenesis	Interference crRNP	Type of nucleic acid targets
1	I	A-G	Cas1, Cas2, Cas4	Cas6/Cas5d	Cascade	DNA
	III	A-F	Cas1, Cas2	Cas6	Csm/Cmr	DNA/RNA
	IV	A-C	Unknown	Csf5	Csf	DNA
2	II	A-C	Cas1, Cas2, Cas4/Csn2, Cas9	RNase III Cas9	Cas9	DNA
	V	A-I, K	Cas1, Cas2, Cas4	Cas12	Cas12	DNA
	VI	A-D	Cas1, Cas2	Cas13	Cas13	RNA



Intensive investigations on model type I CRISPR-Cas systems have revealed molecular mechanisms of multi-Cas CRISPR-based antiviral defense, in respect of crRNA processing, effector complex assembly, PAM recognition and R-loop formation, and Cas3-executed DNA target destruction (Figures 1C,D).

crRNA Processing

A mature crRNA is an essential element for an active CRISPR-Cas effector complex, in which it functions to guide the recognition of the cognate targets for destruction. It has been reported that biogenesis of mature crRNAs requires cleavage of long CRISPR transcripts within each repeat. In type I systems, an RNA recognition motif (RRM)-containing protein, either a Cas5d for type I-C (Garside et al., 2012; Nam et al., 2012; Punetha et al., 2014) or a Cas6 homolog for rest type I systems (Przybilski et al., 2011; Sashital et al., 2011; Li et al., 2013; Sokolowski et al., 2014; Majumdar et al., 2015; Taylor et al., 2019), recognizes and catalyzes the crRNA maturation. It should be noted that some type III systems also use Cas6 for crRNA processing (Hale et al., 2009; Majumdar et al., 2015; Peng et al., 2015; Nickel et al., 2019; Wei et al., 2019). Structural analyses of different Cas6/Cas5d-RNA complexes revealed that many repeats formed characteristic stable stem-loop structures. Accordingly, Cas6/Cas5d enzymes have evolved distinct mechanisms to overcome the challenges of binding and catalysis of various RNA molecules (Figure 2; Koo et al., 2013; Li, 2014; Shao et al., 2016; Sefcikova et al., 2017).

Mature crRNA production has been reported for each subtype of type I. Repeats of CRISPR arrays from type I-E of *E. coli* and *Thermus thermophilus* and I-F of *Pseudomonas aeruginosa* were reported to form stable canonical stem-loop structures (Figures 2A-C; Haurwitz et al., 2010; Sashital et al., 2011; Sternberg et al., 2012). Interestingly, all of the stems and loops greatly vary in both size and sequences, which serve as bases for specific binding of Cas6 and subsequent cleavage. The Cas6 proteins specifically recognize and cleave the pre-crRNAs in a single turnover fashion, forming a stable hairpin structure in each repeat (Haurwitz et al., 2010; Sashital et al., 2011; Sternberg et al., 2012). After cleavage, they remained binding to the cleavage products by associating with the 3' stem-loop structure tightly (Jore et al., 2011). This is consistent with the finding that Cas6 protein is an integrated part of the respective interference complexes (Brouns et al., 2008; Jore et al., 2011; Wiedenheft et al., 2011a,b; Jackson et al., 2014; Mulepati et al., 2014; Zhao et al., 2014). Similarly, repeats of type I-C from *Bacillus halodurans* also formed stable stem-loop structures, which, however, were produced by a unique Cas5 variant (Garside et al., 2012; Nam et al., 2012; Koo et al., 2013; Punetha et al., 2014) [known as Cas5d where the affix “d” refers to “Dvulg” (Haft et al., 2005)], as a *cas6* gene is missing from this system. A 3-bp stem and a tetraloop included in the hairpin region were illuminated to be a minimal structural requirement for Cas5d recognition and cleavage (Nam et al., 2012).

Differently, CRISPR repeats present in type I-A of *Sulfolobus solfataricus* and type I-B of *Methanococcus maripaludis* usually form mini stem-loop structures (Figures 2D,E), instead of showing clear palindromic features (Kunin et al., 2007;

Shao et al., 2016). For instance, in type I-A system, the formed stem-loop is composed of a 3-bp stem and a 5-nt loop, which were specifically recognized and stabilized by SsCas6, a Cas6 homolog. Two base pairs at the base of the stem were essentially required for SsCas6 binding and cleavage, as mutations disrupting the base-pair matching resulted in cleavage inhibition (Shao and Li, 2013). Likewise, a 3-bp stem could be potentially formed by the CRISPR repeats from another type I-B system of *Haloferax volcanii* (Maier et al., 2013), which likely contributes to the binding specificity of the HVCas6. Comparably, type I-G system shares the similar Cas6 homolog protein with coexisting types I-A and III-B systems in *Pyrococcus furiosus* to recognize CRISPR repeats with no obvious stem-loop structure for mature crRNA production (Figure 2F; Carte et al., 2010; Elmore et al., 2015).

crRNA processing in a type I-D system is the least known. Processing of CRISPR transcripts of a type I-D system in *Synechocystis* species PCC 6803 involved a Cas6 protein (Cas6-1) (Scholz et al., 2013; Jesser et al., 2019). Depletion of Cas6-1 from the cells largely decreased the crRNA amount, while this deficient phenotype could be restored through expression of Cas6-1 via a plasmid (Scholz et al., 2013). A mini stem-loop was also predicted to be formed in each repeat (Jesser et al., 2019), but whether it serves as the binding signal for Cas6-1 remains uncertain.

Effector Machinery Assembly

Type I CRISPR-Cas systems encode multiple Cas proteins to multi-Cas effector machineries, which is termed the *CRISPR-associated complex for antiviral defense* (Cascade). Cascade was initially used for characterizing the type I-E CRISPR-Cas effector complex of *E. coli* (Brouns et al., 2008). This complex has a molecular weight of approximately 405 kDa and comprises proteins of Cas8e (also known as CasA or Cse1), Cas11 (also known as CasB or Cse2), Cas7 (also known as CasC or Cse4), Cas5 (also known as CasD), and Cas6e (also known as CasE or Cse3), as well as an RNA component, the 61-nt crRNA (Jore et al., 2011). These proteins shape a seahorse-like architecture with a stoichiometry of Cas8e₁-Cas11₂-Cas7₆-Cas5₁-Cas6e₁ (Figure 1C; Jore et al., 2011).

Two crystal structures of the type I-E Cascade bound to a crRNA have been solved, offering molecular details in the Cascade complex assembly (Figure 3A; Jackson et al., 2014; Zhao et al., 2014). In *E. coli*, as previously mentioned, pre-crRNA is processed by the Cas6e ribonuclease through cleavage of pre-crRNAs within each repeat. After processing, Cas6e remains tight binding to the 3' repeat portion of each crRNA (Jore et al., 2011). At the same time, six Cas7 proteins run along the guide region (spacer) of the crRNA, forming the backbone filament, with the upper most subunit interacting with Cas6e. These Cas7 proteins adopt a right-hand-like shape (Mulepati et al., 2014), and the subunits are connected with one another by interaction between the “thumb” and “fingers” domains of one subunit and its adjacent subunits, respectively (Jackson et al., 2014). The Cas5e protein also exhibits a right-hand-like shape but lacks the “fingers” domain, thus allowing Cas5e to cap the Cas7 filament at the 5' end of the crRNA by undertaking the same interaction of Cas7 with that between Cas7 subunits. In addition, Cas5e protein also holds the first 6 nt of the 5' repeat portion of the crRNA

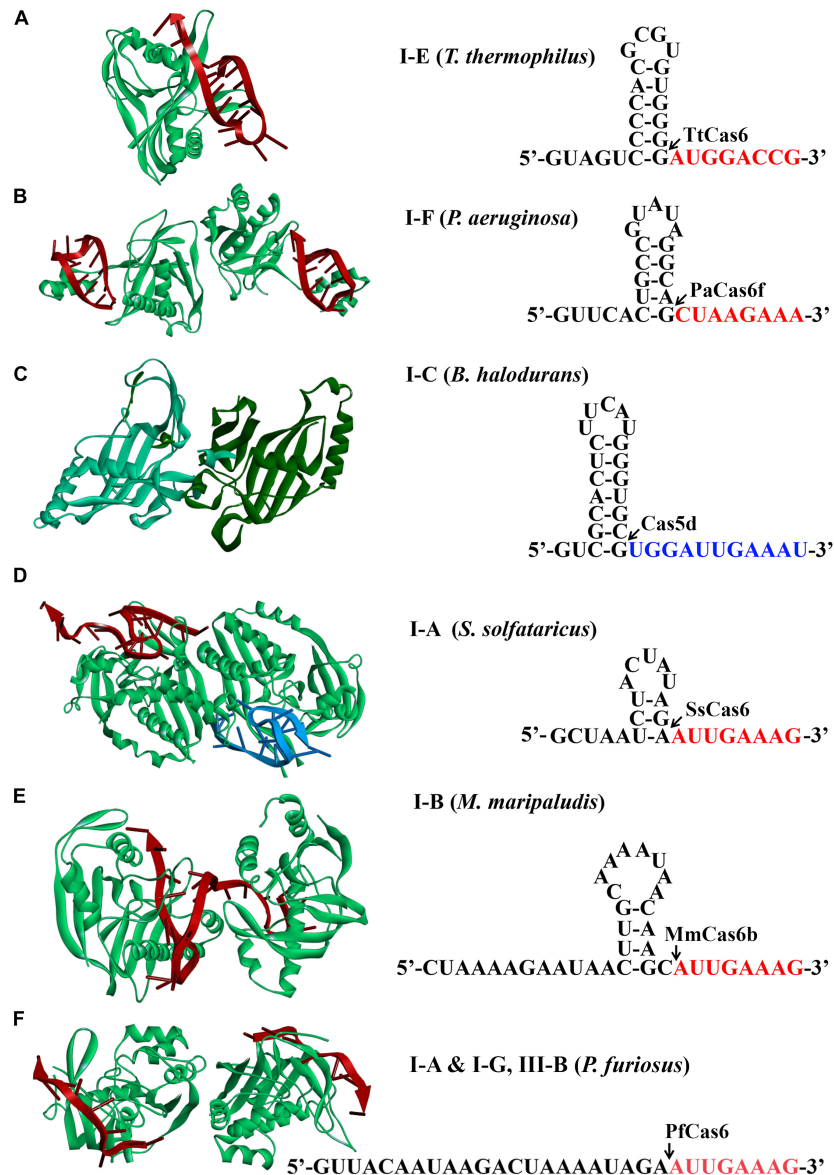
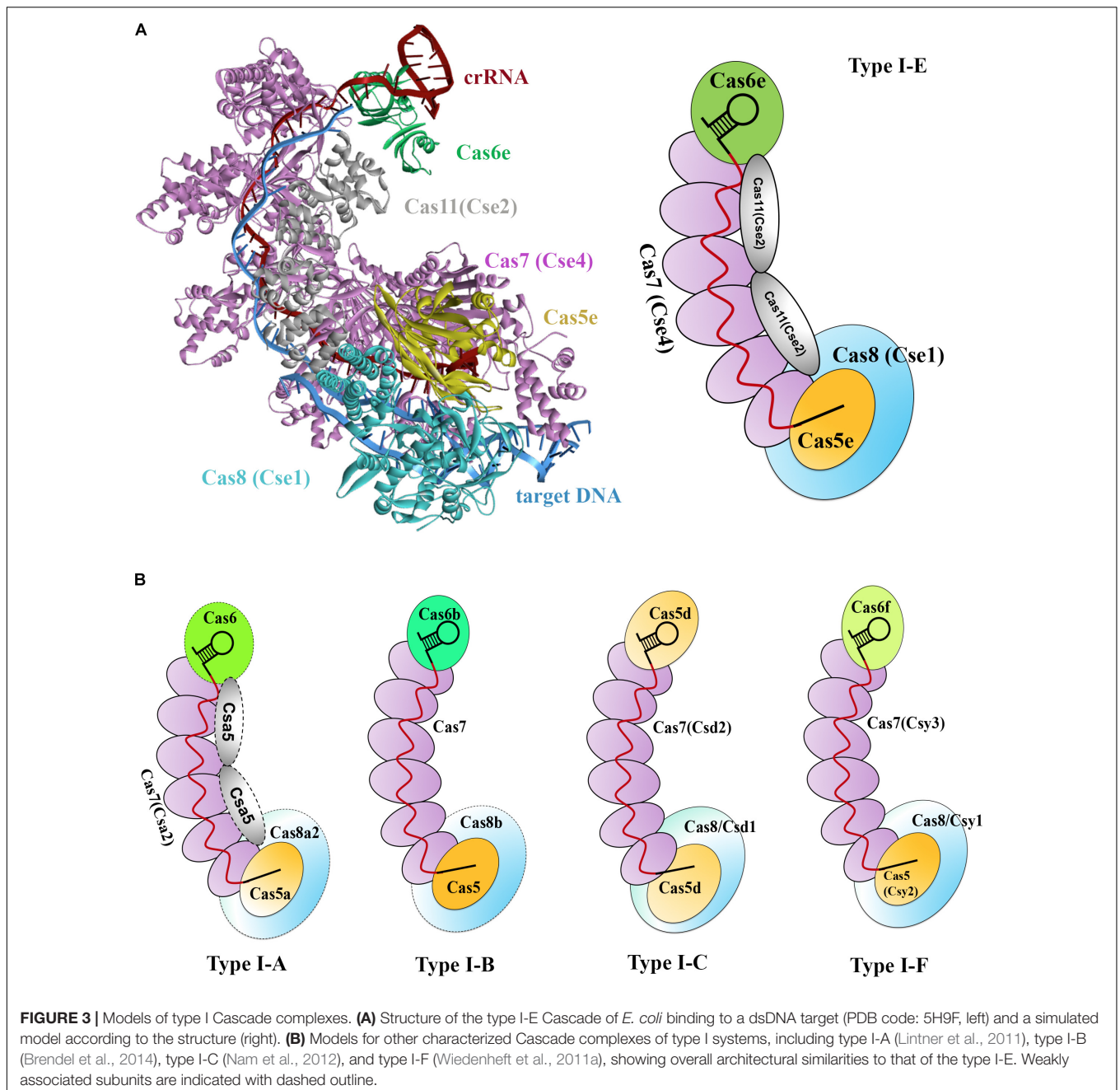


FIGURE 2 | Structures of Cas6 and Cas5d (left) and schematic of CRISPR repeats (right). Structural examples for Cas6 or Cas5d were presented from (A) type I-E of *Thermus thermophilus* (TtCas6, PDB code 2Y8W) (Sashital et al., 2011), (B) type I-F of *Pseudomonas aeruginosa* (PaCas6, PDB code 2XLK) (Haurwitz et al., 2010), (C) type I-C of *Bacillus halodurans* (Cas5d, PDB code 4F3M) (Nam et al., 2012), (D) type I-A of *Sulfolobus solfataricus* (SsCas6, PDB code 4ILL) (Shao and Li, 2013), (E) type I-B of *Methanococcus maripaludis* (MmCas6, PDB code 4Z7K) (Shao et al., 2016), and (F) type I-G of *Pyrococcus furiosus* (PfCas6, PDB code 3PKM) (Wang et al., 2011). The structures of Cas6 are shown in complex with crRNA (red). CRISPR RNAs from type I systems form either stable canonical stem-loop (A–C), mini-stem-loop (D,E), or non-stem-loop (F), structures. The predicted processing sites are indicated by arrows, and sequences are to be present as the 5' handle of a crRNA is shown in red or blue fonts.

(Jackson et al., 2014). In the seahorse-like complex, two Cas11 proteins form a dimer that does not directly interact with crRNA but functions as a bridge to connect the head and tail of the complex. One Cas11 subunit that is proximal to the 5' end of crRNA, along with Cas5e, contacts Cas8 at the bottom of the complex, whereas the other Cas11 subunit interacts with Cas6e on the top (Jackson et al., 2014). Cas8 and Cas11 were identified as the largest and smallest subunits of the Cascade, respectively (Makarova et al., 2011a).

Subsequently, several other type I systems complexes were isolated and characterized (Lintner et al., 2011; Nam et al., 2012; Brendel et al., 2014). Comparison of the overall architectures of these complexes with that of the type I-E Cascade revealed that they share striking architectural similarities and thus being all referred as Cascade (Figure 3B; Reeks et al., 2013). These Cascade complexes show some common features. (1) A core complex of Cas5, Cas7, and/or Cas6 holds a crRNA, each protein possessing at least one RRM motif. (2) Cas7 subunits form the



backbone of the complex and are more abundantly present. (3) The large subunit (Cas8) and/or small subunit (Cas11), if present, less tightly associate with the core complex. Thus far, a separate small subunit is only seen in type I-A and I-E systems, that is, Csa5 and Cse2, respectively. This small subunit is absent from the Cascade complexes of type I-B, I-C, and I-F, but the large subunits of these complexes are speculated to have contained a domain that is functionally homologous to the small subunit (Makarova et al., 2011a). Interestingly, the determined Cascade structure of a type I-F variant (I-Fv) lacks both the large and the small subunits, whose functions, however, are replaced by Cas5fv and Cas7fv, respectively (Pausch et al., 2017). The architectural

similarities suggest that these Cascade complexes may use similar mechanisms for complex assembly and DNA interference.

PAM Recognition and R-Loop Formation

The Cascade is directed to an invading DNA molecule solely relying on base pairing between the embedded crRNA and protospacer. An outstanding question would be how crRNA differentiates protospacer from the corresponding spacer that is stored in the genomic CRISPR bank. The evolved mechanisms for these systems to avoid self-CRISPR targeting involve a sequence immediately flanking the protospacer, called a protospacer adjacent motif (PAM), which is essentially required for Cascade

to determine *bona fide* DNA targets. First predicted by Mojica et al. (2009) and experimentally demonstrated for type I-A systems in *Sulfolobus* by Gudbergstottir et al. (2011), PAM is typically of two to five base pairs and located at the 5' end of protospacer on the strand matching the spacer.

Protospacer adjacent motif recognition has also been studied in other type I systems. In type I-E system of *E. coli*, a loop structure, named the L1 loop that is located within the N-terminal domain of Cas8, directly contacts with the PAM, mediating specific binding of Cascade to the PAM-containing DNA target (Sashital et al., 2012). Further analyses defined that three structural features, including a glutamine wedge, a glycine loop, and a lysine finger, were required for PAM recognition by Cas8 and specified the interaction of PAMs with the target strand (Hayes et al., 2016). In type I-E Cascade of *Thermobifida fusca*, the Cas8 subunit played the same roles in specifying the PAMs, while through contacting the non-target strand (Xiao et al., 2017). In both cases, PAM sequences were recognized by the Cas8 proteins at the minor groove side, explaining the promiscuity of PAM recognition in these systems. Additionally, the recognition of PAM by Cas8 homologs was also identified in type I-B systems of *H. volcanii* and *Methanothermobacter thermautotrophicus* (Cass et al., 2015). Interestingly, in the reported type I-Fv system, the Cas8 protein is missing. Instead, the existing Cas5f variant (Cas5fv) containing an additional domain may compensate for the roles of Cas8 (Pausch et al., 2017).

Bona fide PAM recognition and base pairing between a crRNA and the cognate protospacer determine a target DNA. Interestingly, in some initial work, researchers found that a few type I CRISPR-Cas systems exhibited tolerance to mismatches between the crRNAs and corresponding protospacers. Impacts of mismatches at different regions of the crRNA on target DNA interference markedly varied. In one of the studies, Wiedenheft et al. (2011b) used isothermal titration calorimetry to investigate DNA-binding affinity of the *P. aeruginosa* type I-F system. The results indicated that an 8-nt ssDNA oligo matching the first 8-nt 5' guide sequence of a crRNA (1–8 nt) showed a high binding affinity of the crRNA, whereas another 8-nt ssDNA oligo complementary to the corresponding 5 to 12 region of the crRNA presented a 4-fold weaker binding affinity. Strikingly, all other tested 8-nt ssDNA oligos that matched the crRNA at a region outside the 1- to 8-nt stretch exhibited no measurable binding affinity. These results indicated that the 1- to 8-nt sequence within the crRNA, immediately adjacent to the PAM, played an essential role in defining an invading DNA as an attacker. This is in analogy to the seed sequence in small RNAs, which functions in target recognition in the RNAi in eukaryotes (Lewis et al., 2003).

Following binding of Cascade to a DNA target upon PAM recognition is the formation of a full R-loop, which was observed in several type I systems. Cascade binding could destabilize the target DNA duplex, allowing crRNA to first pair with the protospacer within the seed region and then throughout the whole matching sequences and thus forming a full R-loop, where the non-target strand is bound by the Cas11 dimer (Hochstrasser et al., 2014; Szczelkun et al., 2014). The existence of an intermediate seed bubble immediately following the PAM recognition but before the full R-loop formation was

recently evidenced. Xiao et al. (2017) captured a structural snapshot of the *T. fusca* type I-E Cascade, showing an 11-bp unwound sequence in the seed region. R-loop formation is accompanied by conformational changes of the small and large subunits that trigger recruitment of Cas3 for target degradation (Hochstrasser et al., 2014).

DNA Target Degradation

As aforementioned, Cascade recruited Cas3 to unwind, cut, and degrade the targets. Cas3 typically comprises an N-terminal HD phosphohydrolase domain and a C-terminal superfamily 2 helicase domain (Makarova et al., 2006). There are also exceptional cases in some type I systems. For example, these two domains are encoded as individual proteins, Cas3' (helicase) and Cas3'' (nuclease), respectively, in type I-A of *S. islandicus* Rey15A (Figure 1B; Guo et al., 2011; Makarova et al., 2011b). It was speculated that the HD domain carried out the divalent ion-dependent catalytic cleavage on ssDNA and/or RNA (Han and Krauss, 2009; Beloglazova et al., 2011; Mulepati and Bailey, 2011; Sinkunas et al., 2011), whereas the helicase domain unwound the DNA/DNA and/or DNA/RNA duplexes in the presence of ATP (Howard et al., 2011; Sinkunas et al., 2011).

The mechanism of target DNA degradation by Cas3 was revealed by structural analyses of type I-E targeting complexes, in which the recruitment of Cas3 strictly required the formation of a full R-loop structure. The ape Cascade and R-loop-forming Cascade present as different conformers and Cas3 only selectively capture the latter (Xiao et al., 2017). This conformational difference might act as signals for triggering Cas3 recruitment and also help the type I systems avoid cleaving partially complementary sequences (off-targeting). Upon the formation of Cas3-Cascade-DNA target complex (Huo et al., 2014), Cas3 nuclease nicked the non-target strand in the 7- to 11-nt region of a protospacer (Hochstrasser et al., 2014). The non-target strand contained a significant bulge structure in the R-loop that was created by the Cascade complex, facilitating its handover from Cascade to the bound Cas3 for nicking (Xiao et al., 2017). This handover was recently confirmed to be essential for type I-E immunity (Xiao et al., 2018). Subsequent exonucleolytic degradation of the DNA target was observed to occur in the direction of 3' to 5' with the need of ATP. Sinkunas et al. (2013) found that Cas3 cleaved only on the displaced strand in the R-loop in absence of ATP, but destroyed both strands when ATP was present. This may indicate that unwinding DNA target by the ATP-dependent Cas3 helicase domain could further provide ssDNA substrates for the nuclease domain of Cas3, or for other host nucleases, leading to degradation of the entire DNA target (Brouns et al., 2008; Mulepati and Bailey, 2013).

A Deduced General Mechanism of Type I DNA Interference

Collectively, as exemplified for the type I-E system in Figure 4A, all known type I CRISPR-Cas systems form a Cascade machinery to perform DNA interference in several steps, consisting of multiple protein components complexed with a mature crRNA molecule. A general mechanism of type I Cascade-mediated DNA

interference could be deduced as described following and shown in **Figure 4B**. The interference begins with the recognition of a suitable PAM by Cas8. Direct interaction between the L1 loop of Cas8 and the PAM (Hochstrasser et al., 2014) locates the Cascade to the DNA target and destabilizes the DNA duplex (Szczelkun et al., 2014). The crRNA base pairs with the protospacer, first within the seed region and then throughout the whole matching sequences, to eventually displace the non-target strand, forming a full R-loop. On the other hand, the initial dsDNA binding induces a major conformational change to the Cascade (Wiedenheft et al., 2011a), which may trigger the recruitment of Cas3 to the Cas8 docking site (Westra et al., 2012; Hochstrasser et al., 2014). It was reported that once the non-target strand is displaced by the crRNA, it is exposed and handed over to the Cas3 nuclease for nicking and successive degrading in the 3' to 5' direction (Sinkunas et al., 2013; Gong et al., 2014). This reaction may generate an intermediate degradation product, as partially ssDNA might not result in complete target DNA degradation. In fact, complete degradation of the target DNA is probably mediated either by the nuclease domain of Cas3 or by other host nucleases with assistance of the ATP-dependent helicase domain of Cas3 to unwind the DNA target (Brouns et al., 2008; Mulepati and Bailey, 2013).

PRACTICAL APPLICATIONS OF TYPE I CRISPR-CAS SYSTEMS

The above mechanistic dissections of type I CRISPR-Cas-based immunity have provided solid theoretical basis for exploiting these systems for a wide range of practical applications in bacterial and archaeal hosts. By simply designing a plasmid-borne artificial mini-CRISPR to express genome-targeting crRNAs, a type I Cascade complex can be readily directed to a genome sequence to accomplish various tasks (**Figure 5A**). These applications can be briefly classified into two classes according to the different activities of these systems. One utilizes the intact DNA interference function of Cascade and Cas3 (**Figure 5B**), for example, genome editing, natural variants selection, and antimicrobials, and the other uses only the target surveillance and binding ability of Cascade alone (without Cas3 nuclease) (**Figure 5C**), for example, gene expression regulation.

Genome Editing

To date, genome editing predominates over other CRISPR applications and remains the best developed, however, overwhelmingly focusing on eukaryotes. In the eukaryotic organisms, double-stranded DNA breaks (DSDBs) introduced by CRISPR-Cas systems at specific locations of the genome could be repaired through either the cellular non-homologous end joining (NHEJ) or the homologous recombination (HR) repair pathways. The NHEJ repair system functions in an error-prone manner that usually generates indels (insertions/deletions) within the vicinity of the target site, leading to frame shifts or gene disruption. In contrast, HR repair system normally creates precise changes at desired positions in the genome when providing engineered DNA templates. Dissimilarly, most

prokaryotic organisms harbor only the HR system. It has to be pointed out that although a bacterial version of NHEJ, involving two proteins, Ku and LigD, was reported (Weller et al., 2002), it occurs only in a very small portion of bacteria. In general, prokaryotes repair the CRISPR-caused DSDBs with relatively much lower efficiencies, which is consistent with the frequently observed CRISPR-mediated lethality when self-targeting crRNAs of active systems were introduced into the prokaryotic cells (Edgar and Qimron, 2010; Vercoe et al., 2013; Gomaa et al., 2014; Peng et al., 2015). It is reasonable because the invaders would persist to affect the host cells if the cleavage of invading attackers, such as bacteriophages and conjugative plasmids, by Cas nucleases is subsequently repaired. In practical applications, this is the main obstacle that needs to be addressed in harnessing the endogenous type I systems for genome editing.

Via coupling the DNA targeting activity of endogenous type I CRISPR-Cas systems with the DNA-assisted HR mechanism, various genome editing purposes can be achieved, including gene insertion, deletion, replacement, tagging, and nucleotide substitution, as well as multiplexed gene deletion, in their natural hosts. Here, we take our recently published work on type I-F CRISPR-Cas system as an example to introduce the procedure, strategy, and principle of exploiting an endogenous type I system for genome editing in *Zymomonas mobilis* (Zheng et al., 2019). In the study, we first demonstrated the DNA cleavage activity of the endogenous type I-F CRISPR-Cas of *Z. mobilis* and determined the PAM sequences essential for the system through a plasmid invader assay (Peng et al., 2013, 2015). This activity was then directed to genomic locations by crRNAs produced from plasmid-borne artificial CRISPR loci to generate DNA injuries. In addition, donor DNAs were supplied as inserts of the same plasmids for facilitating HR, allowing their multiplication via plasmid replication to ensure effective concentration and stability, which has been evidenced to be an efficient approach for enhancing the HR rate and hence genome editing efficiency (Li et al., 2016; Renaud et al., 2016; Cheng et al., 2017). Using this established type I-F CRISPR-Cas platform, various genome engineering purposes were efficiently achieved with efficiencies of up to 100%, including gene deletion, replacement, *in situ* modifications, particularly large genomic fragment deletion, and simultaneous removal of three genes (**Figure 5D**). Moreover, our work demonstrated that depletion of a DNA restriction-modification (R-M) system could also lead to boosted genome engineering efficiency (Zheng et al., 2019). Considering the widespread of R-M systems (Koonin et al., 2017a) and other DNA interference systems (Doron et al., 2018) in microorganisms, the method could be served as an important reference for the development and deployment of similar CRISPR-Cas toolkits in organisms with low efficiency in the wild-type genetic background.

Endogenous type I CRISPR-based genome editing was first reported in *S. islandicus*, and then other organisms. Gene deletion and site-directed mutagenesis were completed in the Crenarchaeota *S. islandicus* by constitutively expressing its native type I-A system (Peng et al., 2013; Li et al., 2016). After transforming a plasmid bearing both the CRISPR target and donor DNA template into the host cells, almost all the

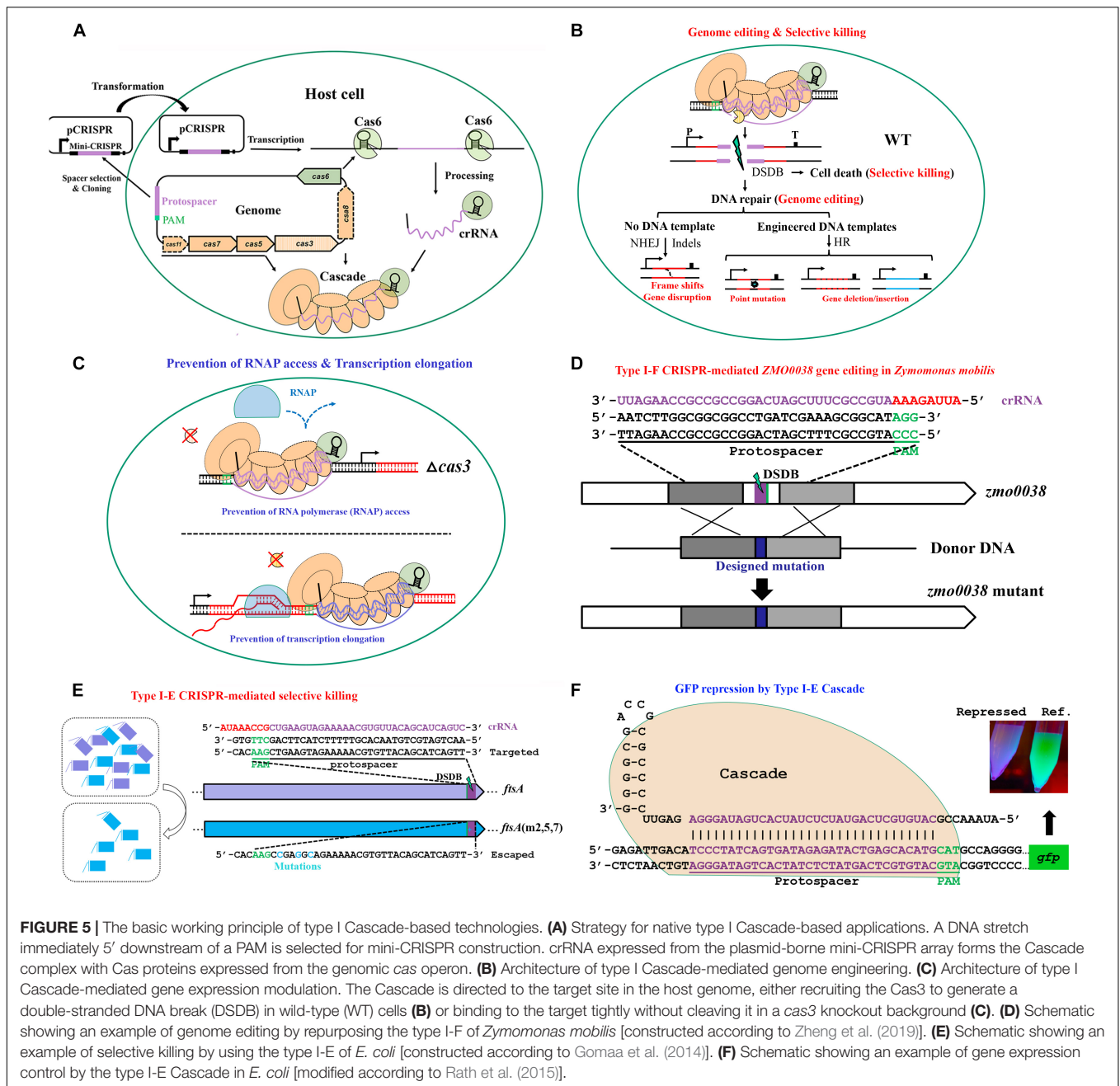


FIGURE 5 | The basic working principle of type I Cascade-based technologies. **(A)** Strategy for native type I Cascade-based applications. A DNA stretch immediately 5' downstream of a PAM is selected for mini-CRISPR construction. crRNA expressed from the plasmid-borne mini-CRISPR array forms the Cascade complex with Cas proteins expressed from the genomic cas operon. **(B)** Architecture of type I Cascade-mediated genome engineering. **(C)** Architecture of type I Cascade-mediated gene expression modulation. The Cascade is directed to the target site in the host genome, either recruiting the Cas3 to generate a double-stranded DNA break (DSDB) in wild-type (WT) cells **(B)** or binding to the target tightly without cleaving it in a cas3 knockout background **(C)**. **(D)** Schematic showing an example of genome editing by repurposing the type I-F of *Zymomonas mobilis* [constructed according to Zheng et al. (2019)]. **(E)** Schematic showing an example of selective killing by using the type I-E of *E. coli* [constructed according to Gomaa et al. (2014)]. **(F)** Schematic showing an example of gene expression control by the type I-E Cascade in *E. coli* [modified according to Rath et al. (2015)].

analyzed transformants were found to possess precise changes including deletions and multiple point substitutions as designed (Li et al., 2016). Likewise, the endogenous type I-B system of the halophilic archaeon *Haloarcula hispanica* was also redirected for gene deletion, insertion, point mutations introduction, and simultaneous deletion of two genes (Cheng et al., 2017). Besides, two other type I-B systems from the bacteria *Clostridium tyrobutyricum* and *Clostridium pasteurianum*, respectively, have also been used for genome editing in their native hosts, with an editing efficiency of 100% on gene deletion/insertion (Pyne et al., 2016; Zhang et al., 2018). Strikingly, the native type I-B system of *C. pasteurianum* facilitated more efficient genome

editing of around 4-fold higher than that of the heterologously expressed Cas9 system (Pyne et al., 2016). Very recently, the type I-E system of *Lactobacillus crispatus* was also exploited to perform *in situ* genomic modifications in the host cells (Hidalgo-Cantabrana et al., 2019b).

Targeted DNA integration can also be accomplished using crRNA-guided transposition, which was first proposed by Peters et al. (2017). It was found that a number of CRISPR-Cas systems are carried by transposons belonging to the Tn7 family (Peters, 2014), and they possess a notable feature of lacking a key factor responsible for DNA targeting. Moreover, the CRISPR-Cas systems carried by them are not capable for DNA interference.

This proposed mechanism has been recently experimentally confirmed in *E. coli* by two groups (Klompe et al., 2019; Strecker et al., 2019). In one instance, Klompe et al. (2019) show that a variant type I-F system from *Vibrio cholerae*, where the adaptation module and Cas3 nuclease are missing, could still form a Cascade complex and guide transposition into specific sites 46 to 55 bp downstream of crRNA-matching sequences. Interestingly, the transposition can be done only with the coevolved type I-F variant, as other tested native CRISPR-Cas systems did not work properly (Klompe et al., 2019). Moreover, the authors also found that the Cascade could be copurified with a TniQ protein, an element of the Tn7 system involved in one of the characterized Tn7 transposition pathways (Peters, 2014). More interestingly, this crRNA-guided transposition could be observed only when expressing a fusion of TniQ-Cas6, but not TniQ-Cas8. These observations thus allowed the authors to propose a mechanism that the CRISPR-Cas system could cofunction with the Tn7 system (Klompe et al., 2019). Similarly, Strecker et al. (2019) reported that the Tn7-like system could recruit a type V-K variant from *Scytonema hofmannii*. This mechanism was later further supported by a cryo-EM structure of TniQ-Cascade complex (Halpin-Healy et al., 2019), within which two TniQ subunits form a dimer, with each interacting with Cas6 and the immediately adjacent Cas7 subunit, respectively. Finally, with either system, CRISPR-directed DNA insertion has been attained in a programmable fashion with high frequencies. More importantly, with no requirement for target DNA cleavage, homologous DNA templates, endogenous DNA repair systems, and selective pressure, this strategy is advantageous for genome modification over many other existing tools.

Selective Killing

As aforementioned, reprogramming an active Cascade toward the chromosome without providing a repair template would lead to killing of most, if not all, targeted cells. Two alternative fates might be brought to the targeted cells, including the death of wild-type cells due to CRISPR-mediated chromosomal degradation, or the survival if variants carry mutations in the targeted sequences. This selective killing feature can be adopted to exploit DNA-cleaving CRISPR-Cas systems as screening tools to select for genetic variants from a given population or as novel programmable antimicrobials to selectively remove certain pathogens from a mixed population (Figure 5B).

In an early study, Edgar and Qimron (2010) demonstrated that the native type I-E CRISPR-Cas system of *E. coli* could be used to cure the cells from prophage. The induced endogenous system could lead to killing of more than 98% of the cells in the population when the targeting activity was directed to an integrated lambda prophage. Interestingly, although prophage induction was also lethal to the cells, simultaneously inducing both of the pathways largely enhanced the survival rate. Further analyses showed that the genome of survivors exclusively did not harbor any prophage sequences. According to the observations, the authors suggested that the survived cells were possibly protected by the CRISPR system acting on excised bacteriophages, and most cells that still harbored prophage genes in the genome were killed by DNA targeting

(Edgar and Qimron, 2010). Later, self-targeting by a native type I-F system was reported to result in dramatic changes in the host's genome (Vercoe et al., 2013). Either part of or an entire preexisting pathogenicity island was deleted from the chromosome when targeted by the type I-F CRISPR-Cas system of *Pectobacterium atrosepticum*, a potato phytopathogen. The chromosomal alterations allowed the genetic variants to survive from CRISPR targeting and resulted in strains lacking pathogenicity, which was proposed to be a strong selective pressure for bacterial evolution (Vercoe et al., 2013).

The sequence specificity and selective killing feature of CRISPR-mediated genome targeting have enabled its antimicrobial utilization (Figure 5E), a concept that was put forward by Gomaa et al. (2014). Gomaa and coworkers assessed the genome targeting efficiency of the native type I-E CRISPR-Cas system from *E. coli* through evaluating cell escape rates. It was found that targeting single or multiple sites within coding or non-coding region of essential or non-essential genes on either strand of the genome gave similar outcomes of a dramatically low cell escape rate, demonstrating that the targeting was efficient and specific yet flexible, requiring only the target sequence with an optimal PAM. This specificity and flexibility allowed selective removal of individual or multiple strains, including highly similar ones, from pure or mixed cultures by directing the targeting activity to a unique or shared PAM-flanking genomic sequence (Gomaa et al., 2014). The application of type I-E CRISPR-Cas system of *E. coli* as antimicrobials was echoed in another study, in which Yosef et al. (2015) engineered a lambda prophage to deliver an active CRISPR-Cas system along with an artificial CRISPR that expresses crRNAs targeting antibiotic resistance genes into the *E. coli* host cells. After the delivery, the lysogenized *E. coli* cells were immunized against the corresponding antibiotic resistance gene-housing genetic elements, not only clearing the preexisting resistance determinants to sensitize the antibiotic-resistant bacteria, but also blocking further uptake of the same resistance. Importantly, the sensitized cells were protected by the delivered CRISPR-Cas system against lytic bacteriophages bearing the same targeting sequences, whereas the residual antibiotic-resistant cells would be persistently attacked by such type of bacteriophages, thus bringing out obvious selective benefit to the sensitized cells over the resistant ones. In addition, two endogenous type II CRISPR-Cas systems of *Streptococcus thermophilus* were also proven to be able to selectively remove closely related organisms by targeting unique sequences from a mixed population of microbes (Gomaa et al., 2014). Therefore, potent selective killing can be gained through different endogenous CRISPR-Cas systems. In fact, if provided with an appropriate delivery vehicle, heterologous CRISPR-Cas can be used as a programmable antimicrobial as well. In two parallel works, a heterologous type II CRISPR-Cas system was delivered into *E. coli* by an M13-based phagemid (Citorik et al., 2014) and into *Staphylococcus aureus* by a Φ NM1-based phagemid (Bikard et al., 2014), respectively, to selectively killing the pathogenic host cells.

In all above cases, both types I and II CRISPR self-targeting could give very robust selective pressure, thus obscuring the difference of targeting effect between these two types of

CRISPR-Cas. As illustrated in **Figure 4B**, Cas3 in type I is mainly responsible for target destruction, first nicking and then degrading DNA via its 3'- to -5'' exonuclease activity. In contrast, Cas9 in type II only cuts the target DNA. It would be interesting to evaluate whether DNA degradation contributes to further increased potency of killing through DNA repair prevention. Nevertheless, these CRISPR-based antimicrobials will undoubtedly limit the increasing threat of antimicrobial resistance microorganisms to the global public health.

Gene Expression Modulation

Type I CRISPR-Cas systems use the protein elements of Cascade and Cas3 for DNA interference. Depletion of Cas3 from a native type I system would still allow Cascade to tightly bind to the target DNA but without cleaving it. Thereby, the standalone Cascade becomes a transcriptional barrier to prevent either RNA polymerase access or transcription elongation, thus being capable of gene expression regulation (**Figure 5C**).

The capability of type I CRISPR-Cas systems in transcriptional regulation has been demonstrated for type I-E system of *E. coli* (**Figure 5F**) in two parallel studies, in both of which the Cas3 nuclease was eliminated (Luo et al., 2015; Rath et al., 2015). In *E. coli*, expression of the Cascade operon is largely repressed under laboratory conditions (Pul et al., 2010; Westra et al., 2010). In order to address this issue, Luo et al. (2015) generated a mutant strain, in which the transcription of Cascade operon was modified to be driven by a constitutive promoter, and Rath et al. (2015) exogenously introduced an extra copy of Cascade operon via a plasmid. These strains kept the ability to process the plasmid-borne CRISPR arrays containing either a single spacer or multiple spacers. As expected, the produced crRNA guided the Cascade to the target sites and exhibited strong repression effect on expression of a genome-integrated *gfp* reporter gene, especially when targeting the promoter region. Moreover, simultaneously repressing several genes was also assayed to be effective by taking the advantage that multiple unique mature crRNA molecules could be produced from a single CRISPR array (Luo et al., 2015; Rath et al., 2015).

Stachler and Marchfelder (2016) similarly, used the endogenous type I-B CRISPR-Cas system to perform *in vivo* gene down-regulation in the archaeon *H. volcanii*. Also, the *cas3* gene was either deleted or mutated to disable its catalytic activity to avoid genome cleavage. In addition, the crRNA-producing Cas6 protein was further depleted from the system, such that the genome-encoded CRISPR arrays would not be processed. Instead, the targeting crRNAs were encoded from a plasmid (Stachler and Marchfelder, 2016) and processed into mature crRNAs via a Cas6-independent approach for *H. volcanii* (Maier et al., 2015). This was evidenced to be significant for the interference efficiency, suggestive of competition between the targeting crRNA and internal crRNAs for Cascade and pointing toward the fact that Cascade is a limiting factor of type I silencing in this strain. Based on the strain modifications, efficient gene repression was seen, with the greatest reducing of transcripts by 92% when targeting the promoter region. Strand bias was also observed that Cascade targeting to the template strand showed much higher silencing efficiency than targeting

to the non-template strand (Stachler and Marchfelder, 2016). In our recent work, gene expression repression by an endogenous type I-F system was also successfully achieved for *Z. mobilis* (Zheng et al., 2019).

Aside from repression of independent genes, type I CRISPR interference was coopted for pathway engineering as well. In a pioneering work, the endogenous type I-E system was used to regulate catabolism of four sugars by targeting their respective operons in *E. coli*. All the four targeted sugar catabolism pathways were efficiently silenced by simply expressing a four-spacer CRISPR array (Luo et al., 2015). Subsequently, using the same endogenous system targeting *gltA* gene in *E. coli* also highly repressed the citrate synthase, leading to accumulation of acetate and thus regulation of the metabolic flux of central metabolism (Chang et al., 2016). Recently, Tarasava et al. (2018) applied CRISPR interference strategy to simultaneously alter the expression levels of multiple genes associated with 3-hydroxypropionate (3-PH) production and finally achieved nearly 2-fold increase of 3-PH yield in *E. coli*. In order to controllably utilize the type I-E system for CRISPR interference, the authors replaced the native promoter of Cascade operon with an arabinose-inducible promoter in a *cas3* deletion mutant. Considering that the placement of spacers within the cluster would influence the repression strength (Luo et al., 2015), the effect of different orders of spacer combination on the enhancement of 3-PH production was tested. On the other hand, a combinatorial library of gRNA array targeting six genes was built and subsequently used to screen for highest producing variants in this work, which was estimated to contain approximately 10⁴ unique variants and approximately three orders of magnitude more than that analyzed in the rational design approaches (Lv et al., 2015; Wu et al., 2015). Importantly, this library was proven to be diverse combinatorial as 48 out of 50 sequenced variants were found to be unique to each other. Furthermore, the orders of spacer combination harbored by some outstanding producing variants screened from the library were highly consistent with the rationally engineered variants (Tarasava et al., 2018).

All these studies have paved the way for further development of type I CRISPR-Cas-based *in vivo* gene interference tools in bacteria and archaea. Considering the broad occurrence of diverse type I systems, it is of great interest to see whether other subtypes are able to perform gene silencing in many other bacteria and archaea.

CONCLUSION AND PERSPECTIVES

Based on the extensive studies on characterizing the type I CRISPR-Cas functions, a general mechanism of type I DNA interference can be deduced (**Figure 4B**). The functional demonstrations have contributed great efforts for exploiting the endogenous type I systems as CRISPR toolkits for genome editing, antimicrobials, gene regulation, and so on, in bacteria and archaea. While many successful achievements have already proven type I CRISPR-based technologies to be powerful tools for prokaryotic engineering, they are still in their infancy stage,

leaving many hurdles for their deeper improvements for a wider range of applications.

Hurdles Ahead

To truly exploit the endogenous type I CRISPR-Cas systems, there are several critical obstacles still required to be addressed. First, because of the host-specific property of endogenous CRISPR-Cas systems, each of them must be thoroughly characterized in the host cells, including determination of PAM sequences, demonstration of Cas functions, and identification of mature crRNAs, as well as dissection of its mode of action during foreign DNA defense. Second, low efficiency of exogenous DNA delivery and their low stability in host cells might significantly hamper the repurposing of endogenous CRISPR-Cas. Although there is no need of importing exogenous Cas proteins, other genetic materials, for example, artificial CRISPR arrays, and donor DNA templates during genome editing, are still required to be introduced into the cells with relatively high concentrations. Therefore, efficient delivery systems would help for better performances during the applications. Separately, prokaryotes harbor a collection of innate defense systems, such as R-M (Roberts et al., 2015), BREX (Goldfarb et al., 2015), and so on. Many of these systems are prevalent; for example, approximately 95% of sequenced bacterial genomes encode at least one R-M system (Roberts et al., 2015). Research has shown that destroying such innate defense systems can lead to boosted plasmid transformation rate and hence genome engineering efficiency (Zheng et al., 2019). In addition, because most of prokaryotes use the less efficient HR pathway for repairing the CRISPR-generated DNA injuries, it is always good to maintain a high concentration of introduced donor DNAs to ensure a better performance of genome editing. To this end, several strategies can be used, such as supplying DNA templates as inserts of the vectors to allow their multiplication (Li et al., 2016; Cheng et al., 2017; Zheng et al., 2019). Finally, during the CRISPR-targeting, escape always occurs with various modes, normally due to modifications/alterations in spacer or protospacer, or alternatively mutations in Cas proteins. However, it would be important but troublesome to understand the escape modes and underlying mechanisms, which might provide possibility for further increasing the overall efficiency of genome manipulations.

CRISPR-dCas3 Exploitation

Current applications of type I CRISPR-Cas systems focus mainly on genome editing and gene down-regulation. For the latter,

in most studies, the *cas3* gene was deleted from the system. Actually, a system with a catalytically inactive Cas3 (dCas3) would have further uses. For instance, considering DSDBs are repaired with relatively lower efficiencies in prokaryotic cells, it would be beneficial to use a catalytically inactive Cas protein, for example, dCas9 or dCas3, with the *FokI* endonuclease to generate DSDBs in a manner that can be more easily repaired by the endogenous pathways, hence increasing engineering efficiency (Guilinger et al., 2014; Tsai et al., 2014). In an even efficient manner, single DNA changes within a targeted window can be achieved without introducing DSDBs by using deaminase-dCas9 fusions in *E. coli* (Gaudelli et al., 2017; Banno et al., 2018), where the CRISPR-dCas9 system might be potentially replaced with an endogenous CRISPR-dCas3 system. Additionally, dCas3, together with the Cascade, may enrich a fusing transcription activator to a genomic region and may thus lead to activation of targeted genes, just in parallel with the similar applications of dCas9/dCas12 as reviewed by Yao et al. (2018). In short, attempts of other applications through the endogenous type I CRISPR-dCas3 systems, such as cell imaging that has not been reported in prokaryotes, can be achievable and will further extend the application scope of CRISPR-based technologies.

AUTHOR CONTRIBUTIONS

WP conceived the outline with inputs from all authors. YZ and WP prepared and wrote the manuscript. WP and LY revised the manuscript. JL, BW, JH, YH, SW, XM, SY, and LM conducted the extensive review. All authors contributed to the writing, reading, and approval of the final manuscript.

FUNDING

This work was supported by the National Key Technology Research and Development Program of China (2018YFA090091 and 2018YFD0500203 to LY); the Natural Science Foundation of Hubei Province of China (2017CFB538, to WP); the Scientific Research Program of Hubei Provincial Department of Education (Q20161007, to WP); and the National Science Foundation of China (31870057, to LY and WP). WP and LY also acknowledge the support from the State Key Laboratory of Biocatalysis and Enzyme Engineering.

REFERENCES

- Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H., and Kondo, A. (2018). Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* 3, 423–429. doi: 10.1038/s41564-017-0102-6
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140
- Beloglazova, N., Petit, P., Flick, R., Brown, G., Savchenko, A., and Yakunin, A. F. (2011). Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* 30, 4616–4627. doi: 10.1038/emboj.2011.377
- Bikard, D., Euler, C. W., Jiang, W., Nussenzweig, P. M., Goldberg, G. W., Duportet, X., et al. (2014). Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* 32, 1146–1150. doi: 10.1038/nbt.3043
- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151, 2551–2561. doi: 10.1099/mic.0.28048-0
- Brendel, J., Stoll, B., Lange, S. J., Sharma, K., Lenz, C., Stachler, A. E., et al. (2014). A complex of Cas proteins 5, 6, and 7 is required for the biogenesis and stability of clustered regularly interspaced short palindromic repeats (crispr)-derived rnas (crRNAs) in *Haloflex volcanii*. *J. Biol. Chem.* 289, 7164–7177. doi: 10.1074/jbc.M113.508184

- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., et al. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689
- Carte, J., Pfister, N. T., Compton, M. M., Terns, R. M., and Terns, M. P. (2010). Binding and cleavage of CRISPR RNA by Cas6. *RNA* 16, 2181–2188. doi: 10.1261/rna.2230110
- Cass, S. D., Haas, K. A., Stoll, B., Alkhnbashi, O. S., Sharma, K., Urlaub, H., et al. (2015). The role of Cas8 in type I CRISPR interference. *Biosci. Rep.* 35:e00197. doi: 10.1042/BSR20150043
- Chang, Y., Su, T., Qi, Q., and Liang, Q. (2016). Easy regulation of metabolic flux in *Escherichia coli* using an endogenous type I-E CRISPR-Cas system. *Microb. Cell Fact.* 15:195. doi: 10.1186/s12934-016-0594-4
- Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479–1491. doi: 10.1016/j.cell.2013.12.001
- Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., et al. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163–1171. doi: 10.1038/cr.2013.122
- Cheng, F., Gong, L., Zhao, D., Yang, H., Zhou, J., Li, M., et al. (2017). Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon. *J. Genet. Genomics* 44, 541–548. doi: 10.1016/j.jgg.2017.09.010
- Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B., and Cho, B. K. (2018). High-level dCas9 expression induces abnormal cell morphology in *Escherichia coli*. *ACS Synth. Biol.* 7, 1085–1094. doi: 10.1021/acssynbio.7b00462
- Citorik, R. J., Mimee, M., and Lu, T. K. (2014). Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* 32, 1141–1145. doi: 10.1038/nbt.3011
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143
- Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J., et al. (2017). RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027. doi: 10.1126/science.aag0180
- Crowley, V. M., Catching, A., Taylor, H. N., Borges, A. L., Metcalf, J., Bondy-Denomy, J., et al. (2019). A type IV-A CRISPR-Cas system in *Pseudomonas aeruginosa* mediates RNA-guided plasmid interference in vivo. *CRISPR J.* 2, 434–440. doi: 10.1089/crispr.2019.0048
- Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., et al. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359:eaar4120. doi: 10.1126/science.aar4120
- Edgar, R., and Qimron, U. (2010). The *Escherichia coli* CRISPR system protects from lambda lysis, lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192, 6291–6294. doi: 10.1128/JB.00644-10
- Elmore, J., Deighan, T., Westpheling, J., Terns, R. M., and Terns, M. P. (2015). DNA targeting by the type I-G and type I-A CRISPR-Cas systems of *Pyrococcus furiosus*. *Nucleic Acids Res.* 43, 10353–10363. doi: 10.1093/nar/gkv1140
- Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517–521. doi: 10.1038/nature17945
- Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523
- Garside, E. L., Schellenberg, M. J., Gesner, E. M., Bonanno, J. B., Sauder, J. M., Burley, S. K., et al. (2012). Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. *RNA* 18, 2020–2028. doi: 10.1261/rna.033100.112
- Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644
- Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., et al. (2015). BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* 34, 169–183. doi: 10.15252/embj.201489455
- Gomaa, A. A., Klumpe, H. E., Luo, M. L., Selle, K., Barrangou, R., and Beisel, C. L. (2014). Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *mBio* 5:e00928-13. doi: 10.1128/mBio.00928-13
- Gong, B., Shin, M., Sun, J., Jung, C. H., Bolt, E. L., van der Oost, J., et al. (2014). Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16359–16364. doi: 10.1073/pnas.1410806111
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). 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
- Gudbergsdottir, S., Deng, L., Chen, Z., Jensen, J. V., Jensen, L. R., She, Q., et al. (2011). Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* 79, 35–49. doi: 10.1111/j.1365-2958.2010.07452.x
- Guillering, J. P., Thompson, D. B., and Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582. doi: 10.1038/nbt.2909
- Guo, L., Brugger, K., Liu, C., Shah, S. A., Zheng, H., Zhu, Y., et al. (2011). Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. *J. Bacteriol.* 193, 1672–1680. doi: 10.1128/JB.01487-10
- Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1:e60. doi: 10.1371/journal.pcbi.0010060
- Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., et al. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139, 945–956. doi: 10.1016/j.cell.2009.07.040
- Halpin-Healy, T. S., Klompe, S. E., Sternberg, S. H., and Fernandez, I. S. (2019). Structural basis of DNA targeting by a transposon-encoded CRISPR-Cas system. *Nature* 577, 271–274. doi: 10.1038/s41586-019-1849-0
- Han, D., and Krauss, G. (2009). Characterization of the endonuclease SSO2001 from *Sulfolobus solfataricus* P2. *FEBS Lett.* 583, 771–776. doi: 10.1016/j.febslet.2009.01.024
- Harrington, L. B., Burstein, D., Chen, J. S., Paez-Espino, D., Ma, E., Witte, I. P., et al. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* 362, 839–842. doi: 10.1126/science.aav4294
- Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010). Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329, 1355–1358. doi: 10.1126/science.1192272
- Hayes, R. P., Xiao, Y., Ding, F., van Erp, P. B., Rajashankar, K., Bailey, S., et al. (2016). Structural basis for promiscuous PAM recognition in type I-E Cascade from *E. coli*. *Nature* 530, 499–503. doi: 10.1038/nature16995
- Hidalgo-Cantabrana, C., Goh, Y. J., and Barrangou, R. (2019a). Characterization and repurposing of type I and type II CRISPR-Cas systems in bacteria. *J. Mol. Biol.* 431, 21–33. doi: 10.1016/j.jmb.2018.09.013
- Hidalgo-Cantabrana, C., Goh, Y. J., Pan, M., Sanozky-Dawes, R., and Barrangou, R. (2019b). Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 15774–15783. doi: 10.1073/pnas.1905421116
- Hochstrasser, M. L., Taylor, D. W., Bhat, P., Guegler, C. K., Sternberg, S. H., Nogales, E., et al. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6618–6623. doi: 10.1073/pnas.1405079111
- Howard, J. A., Delmas, S., Ivancic-Bace, I., and Bolt, E. L. (2011). Helicase dissociation and annealing of RNA-DNA hybrids by *Escherichia coli* Cas3 protein. *Biochem. J.* 439, 85–95. doi: 10.1042/BJ20110901
- Huo, Y., Nam, K. H., Ding, F., Lee, H., Wu, L., Xiao, Y., et al. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat. Struct. Mol. Biol.* 21, 771–777. doi: 10.1038/nsmb.2875
- Jackson, R. N., Golden, S. M., van Erp, P. B., Carter, J., Westra, E. R., Brouns, S. J., et al. (2014). Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*. *Science* 345, 1473–1479. doi: 10.1126/science.1256328
- Jansen, R., Embden, J. D., Gaastra, W., and Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575. doi: 10.1046/j.1365-2958.2002.02839.x
- Jesser, R., Behler, J., Benda, C., Reimann, V., and Hess, W. R. (2019). Biochemical analysis of the Cas6-1 RNA endonuclease associated with the subtype I-D

- CRISPR-Cas system in *Synechocystis* sp. PCC 6803. *RNA Biol.* 16, 481–491. doi: 10.1080/15476286.2018.1447742
- Jiang, W., Yang, B., and Weeks, D. P. (2014). Efficient CRISPR/Cas9-mediated gene editing in *Arabidopsis thaliana* and inheritance of modified genes in the T2 and T3 generations. *PLoS One* 9:e99225. doi: 10.1371/journal.pone.0099225
- Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8:15179. doi: 10.1038/ncomms15179
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829
- Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* 18, 529–536. doi: 10.1038/nsmb.2019
- Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S., and Sternberg, S. H. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* 571, 219–225. doi: 10.1038/s41586-019-1323-z
- Koo, Y., Ka, D., Kim, E. J., Suh, N., and Bae, E. (2013). Conservation and variability in the structure and function of the Cas5d endonuclease in the CRISPR-mediated microbial immune system. *J. Mol. Biol.* 425, 3799–3810. doi: 10.1016/j.jmb.2013.02.032
- Koonin, E. V., Makarova, K. S., and Wolf, Y. I. (2017a). Evolutionary genomics of defense systems in archaea and bacteria. *Annu. Rev. Microbiol.* 71, 233–261. doi: 10.1146/annurev-micro-090816-093830
- Koonin, E. V., Makarova, K. S., and Zhang, F. (2017b). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008
- Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8:R61. doi: 10.1186/gb-2007-8-4-r61
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell* 115, 787–798. doi: 10.1016/s0092-8674(03)01018-3
- Li, H. (2014). Structural principles of CRISPR RNA processing. *Structure* 23, 13–20. doi: 10.1016/j.str.2014.10.006
- Li, M., Liu, H., Han, J., Liu, J., Wang, R., Zhao, D., et al. (2013). Characterization of CRISPR RNA biogenesis and Cas6 cleavage-mediated inhibition of a provirus in the haloarchaeon *Haloferax mediterranei*. *J. Bacteriol.* 195, 867–875. doi: 10.1128/JB.01688-12
- Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., et al. (2016). Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic Acids Res.* 44:e34. doi: 10.1093/nar/gkv1044
- Lintner, N. G., Kerou, M., Brumfield, S. K., Graham, S., Liu, H., Naismith, J. H., et al. (2011). Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* 286, 21643–21656. doi: 10.1074/jbc.M111.238485
- Liu, T., Pan, S., Li, Y., Peng, N., and She, Q. (2018). Type III CRISPR-Cas system: introduction and its application for genetic manipulations. *Curr. Issues Mol. Biol.* 26, 1–14. doi: 10.21775/cimb.026.001
- Luo, M. L., Mullis, A. S., Leenay, R. T., and Beisel, C. L. (2015). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res.* 43, 674–681. doi: 10.1093/nar/gku971
- Lv, L., Ren, Y. L., Chen, J. C., Wu, Q., and Chen, G. Q. (2015). Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes, a case study: controllable P(3HB-co-4HB) biosynthesis. *Metab. Eng.* 29, 160–168. doi: 10.1016/j.ymben.2015.03.013
- Maier, L. K., Lange, S. J., Stoll, B., Haas, K. A., Fischer, S., Fischer, E., et al. (2013). Essential requirements for the detection and degradation of invaders by the *Haloferax volcanii* CRISPR/Cas system I-B. *RNA Biol.* 10, 865–874. doi: 10.4161/rna.24282
- Maier, L. K., Stachler, A. E., Saunders, S. J., Backofen, R., and Marchfelder, A. (2015). An active immune defense with a minimal CRISPR (clustered regularly interspaced short palindromic repeats) RNA and without the Cas6 protein. *J. Biol. Chem.* 290, 4192–4201. doi: 10.1074/jbc.M114.617506
- Majumdar, S., Zhao, P., Pfister, N. T., Compton, M., Olson, S., Glover, C. V., et al. (2015). Three CRISPR-Cas immune effector complexes coexist in *Pyrococcus furiosus*. *RNA* 21, 1147–1158. doi: 10.1261/rna.049130.114
- Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011a). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* 6:38. doi: 10.1186/1745-6150-6-38
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1:7. doi: 10.1186/1745-6150-1-7
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011b). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2019). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. doi: 10.1038/s41579-019-0299-x
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2018). Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J.* 1, 325–336. doi: 10.1089/crispr.2018.0033
- Malina, A., Mills, J. R., Cencic, R., Yan, Y., Fraser, J., Schippers, L. M., et al. (2013). Repurposing CRISPR/Cas9 for in situ functional assays. *Genes Dev.* 27, 2602–2614. doi: 10.1101/gad.227132.113
- Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. *Science* 322, 1843–1845. doi: 10.1126/science.1165771
- Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740. doi: 10.1099/mic.0.023960-0
- Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182. doi: 10.1007/s00239-004-0046-3
- Mulepati, S., and Bailey, S. (2011). Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.* 286, 31896–31903. doi: 10.1074/jbc.M111.270017
- Mulepati, S., and Bailey, S. (2013). In vitro reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J. Biol. Chem.* 288, 22184–22192. doi: 10.1074/jbc.M113.472233
- Mulepati, S., Heroux, A., and Bailey, S. (2014). Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* 345, 1479–1484. doi: 10.1126/science.1256996
- Nam, K. H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M. P., et al. (2012). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* 20, 1574–1584. doi: 10.1016/j.str.2012.06.016
- Nickel, L., Ulbricht, A., Alkhnbashi, O. S., Forstner, K. U., Cassidy, L., Weidenbach, K., et al. (2019). Cross-cleavage activity of Cas6b in crRNA processing of two different CRISPR-Cas systems in *Methanosarcina mazei* Go1. *RNA Biol.* 16, 492–503. doi: 10.1080/15476286.2018.1514234
- O'Connell, M. R. (2019). Molecular mechanisms of RNA targeting by Cas13-containing Type VI CRISPR-Cas systems. *J. Mol. Biol.* 431, 66–87. doi: 10.1016/j.jmb.2018.06.029
- Ozcan, A., Pausch, P., Linden, A., Wulf, A., Schuhle, K., Heider, J., et al. (2019). Type IV CRISPR RNA processing and effector complex formation in *Aromatoleum aromaticum*. *Nat. Microbiol.* 4, 89–96. doi: 10.1038/s41564-018-0274-8
- Pausch, P., Muller-Esparza, H., Gleditsch, D., Altegoer, F., Randau, L., and Bange, G. (2017). Structural variation of Type I-F CRISPR RNA guided DNA surveillance. *Mol. Cell* 67, 622–632.e4. doi: 10.1016/j.molcel.2017.06.036
- Peng, W., Feng, M., Feng, X., Liang, Y. X., and She, Q. (2015). An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res.* 43, 406–417. doi: 10.1093/nar/gku1302

- Peng, W., Li, H., Hallstrom, S., Peng, N., Liang, Y. X., and She, Q. (2013). Genetic determinants of PAM-dependent DNA targeting and pre-crRNA processing in *Sulfolobus islandicus*. *RNA Biol.* 10, 738–748. doi: 10.4161/rna.23798
- Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973–976. doi: 10.1038/nmeth.2600
- Peters, J. E. (2014). Tn7. *Microbiol. Spectr.* 2, 1–20. doi: 10.1128/microbiolspec.MDNA3-0010-2014
- Peters, J. E., Makarova, K. S., Shmakov, S., and Koonin, E. V. (2017). Recruitment of CRISPR-Cas systems by Tn7-like transposons. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7358–E7366. doi: 10.1073/pnas.1709035114
- Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151, 653–663. doi: 10.1099/mic.0.27437-0
- Przybilski, R., Richter, C., Gristwood, T., Clulow, J. S., Vercoe, R. B., and Fineran, P. C. (2011). Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol.* 8, 517–528. doi: 10.4161/rna.8.3.15190
- Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* 75, 1495–1512. doi: 10.1111/j.1365-2958.2010.07073.x
- Punetha, A., Sivathanu, R., and Anand, B. (2014). Active site plasticity enables metal-dependent tuning of Cas5d nuclease activity in CRISPR-Cas type I-C system. *Nucleic Acids Res.* 42, 3846–3856. doi: 10.1093/nar/gkt1335
- Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A., and Chou, C. P. (2016). Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Sci. Rep.* 6:25666. doi: 10.1038/srep25666
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. doi: 10.1016/j.cell.2013.02.022
- Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P. R., and Lundgren, M. (2015). Efficient programmable gene silencing by Cascade. *Nucleic Acids Res.* 43, 237–246. doi: 10.1093/nar/gku1257
- Reeks, J., Naismith, J. H., and White, M. F. (2013). CRISPR interference: a structural perspective. *Biochem. J.* 453, 155–166. doi: 10.1042/BJ20130316
- Renaud, J. B., Boix, C., Charpentier, M., De Cian, A., Cochenne, J., Duvernois-Berthet, E., et al. (2016). Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Rep.* 14, 2263–2272. doi: 10.1016/j.celrep.2016.02.018
- Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2015). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 43, D298–D299. doi: 10.1093/nar/gku1046
- Sashital, D. G., Jinek, M., and Doudna, J. A. (2011). An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* 18, 680–687. doi: 10.1038/nsmb.2043
- Sashital, D. G., Wiedenheft, B., and Doudna, J. A. (2012). Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* 46, 606–615. doi: 10.1016/j.molcel.2012.03.020
- Scholz, I., Lange, S. J., Hein, S., Hess, W. R., and Backofen, R. (2013). CRISPR-Cas systems in the cyanobacterium *Synechocystis* sp. PCC6803 exhibit distinct processing pathways involving at least two Cas6 and a Cmr2 protein. *PLoS One* 8:e56470. doi: 10.1371/journal.pone.0056470
- Sefcikova, J., Roth, M., Yu, G., and Li, H. (2017). Cas6 processes tight and relaxed repeat RNA via multiple mechanisms: a hypothesis. *Bioessays* 39:10.1002/bies.201700019. doi: 10.1002/bies.201700019
- Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., et al. (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. doi: 10.1073/pnas.1104144108
- Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84–87. doi: 10.1126/science.1247005
- Shao, Y., and Li, H. (2013). Recognition and cleavage of a nonstructured CRISPR RNA by its processing endoribonuclease Cas6. *Structure* 21, 385–393. doi: 10.1016/j.str.2013.01.010
- Shao, Y., Richter, H., Sun, S., Sharma, K., Urlaub, H., Randau, L., et al. (2016). A non-stem-loop CRISPR RNA is processed by dual binding Cas6. *Structure* 24, 547–554. doi: 10.1016/j.str.2016.02.009
- Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30, 1335–1342. doi: 10.1038/emboj.2011.41
- Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P., et al. (2013). *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* 32, 385–394. doi: 10.1038/emboj.2012.352
- Sokolowski, R. D., Graham, S., and White, M. F. (2014). Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.* 42, 6532–6541. doi: 10.1093/nar/gku308
- Stachler, A. E., and Marchfelder, A. (2016). Gene repression in Haloarchaea using the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas I-B system. *J. Biol. Chem.* 291, 15226–15242. doi: 10.1074/jbc.M116.724062
- Sternberg, S. H., Haurwitz, R. E., and Doudna, J. A. (2012). Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* 18, 661–672. doi: 10.1261/rna.030882.111
- Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J. L., Makarova, K. S., Koonin, E. V., et al. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365, 48–53. doi: 10.1126/science.aax9181
- Szczelkun, M. D., Tikhomirova, M. S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., et al. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9798–9803. doi: 10.1073/pnas.1402597111
- Takei, Y., Shah, S., Harvey, S., Qi, L. S., and Cai, L. (2017). Multiplexed Dynamic imaging of genomic loci by combined CRISPR imaging and DNA sequential FISH. *Biophys. J.* 112, 1773–1776. doi: 10.1016/j.bpj.2017.03.024
- Tarasava, K., Liu, R., Garst, A., and Gill, R. T. (2018). Combinatorial pathway engineering using type I-E CRISPR interference. *Biotechnol. Bioeng.* 115, 1878–1883. doi: 10.1002/bit.26589
- Taylor, H. N., Warner, E. E., Armbrust, M. J., Crowley, V. M., Olsen, K. J., and Jackson, R. N. (2019). Structural basis of Type IV CRISPR RNA biogenesis by a Cas6 endoribonuclease. *RNA Biol.* 16, 1438–1447. doi: 10.1080/15476286.2019.1634965
- Tsai, S. Q., Wuyekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., et al. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576. doi: 10.1038/nbt.2908
- Vercoe, R. B., Chang, J. T., Dy, R. L., Taylor, C., Gristwood, T., Clulow, J. S., et al. (2013). Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet.* 9:e1003454. doi: 10.1371/journal.pgen.1003454
- Wang, R., Preamplume, G., Terns, M. P., Terns, R. M., and Li, H. (2011). Interaction of the Cas6 ribonuclease with CRISPR RNAs: recognition and cleavage. *Structure* 19, 257–264. doi: 10.1016/j.str.2010.11.014
- Wei, W., Zhang, S., Fleming, J., Chen, Y., Li, Z., Fan, S., et al. (2019). Mycobacterium tuberculosis type III-A CRISPR/Cas system crRNA and its maturation have atypical features. *FASEB J.* 33, 1496–1509. doi: 10.1096/fj.201800557RR
- Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., et al. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686–1689. doi: 10.1126/science.1074584
- Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., et al. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77, 1380–1393. doi: 10.1111/j.1365-2958.2010.07315.x
- Westra, E. R., van Erp, P. B., Kunne, T., Wong, S. P., Staals, R. H., Seegers, C. L., et al. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46, 595–605. doi: 10.1016/j.molcel.2012.03.018

- Wiedenheft, B., Lander, G. C., Zhou, K., Jore, M. M., Brouns, S. J., van der Oost, J., et al. (2011a). Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 477, 486–489. doi: 10.1038/nature10402
- Wiedenheft, B., van Duijn, E., Bultema, J. B., Waghmare, S. P., Zhou, K., Barendregt, A., et al. (2011b). 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. doi: 10.1073/pnas.1102716108
- Wu, J., Du, G., Chen, J., and Zhou, J. (2015). Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Sci. Rep.* 5:13477. doi: 10.1038/srep13477
- Xiao, Y., Luo, M., Dolan, A. E., Liao, M., and Ke, A. (2018). Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science* 361:eaat0839. doi: 10.1126/science.aat0839
- Xiao, Y., Luo, M., Hayes, R. P., Kim, J., Ng, S., Ding, F., et al. (2017). Structure basis for directional R-loop formation and substrate handover mechanisms in Type I CRISPR-Cas system. *Cell* 170, 48–60.e11. doi: 10.1016/j.cell.2017.06.012
- Yao, R., Liu, D., Jia, X., Zheng, Y., Liu, W., and Xiao, Y. (2018). CRISPR-Cas9/Cas12a biotechnology and application in bacteria. *Synth. Syst. Biotechnol.* 3, 135–149. doi: 10.1016/j.synbio.2018.09.004
- Yosef, I., Manor, M., Kiro, R., and Qimron, U. (2015). Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7267–7272. doi: 10.1073/pnas.1500107112
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038
- Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737
- Zhang, J., Zong, W., Hong, W., Zhang, Z. T., and Wang, Y. (2018). Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab. Eng.* 47, 49–59. doi: 10.1016/j.ymben.2018.03.007
- Zhao, H., Sheng, G., Wang, J., Wang, M., Bunkoczi, G., Gong, W., et al. (2014). Crystal structure of the RNA-guided immune surveillance Cascade complex in *Escherichia coli*. *Nature* 515, 147–150. doi: 10.1038/nature13733
- Zheng, Y., Han, J., Wang, B., Hu, X., Li, R., Shen, W., et al. (2019). Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* 47, 11461–11475. doi: 10.1093/nar/gkz940

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zheng, Li, Wang, Han, Hao, Wang, Ma, Yang, Ma, Yi and Peng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.