Characterization and Repurposing of Type I and Type II CRISPR–Cas Systems in Bacteria

Claudio Hidalgo-Cantabrana, Yong Jun Goh and Rodolphe Barrangou

Department of Food, *Processing and Nutrition Sciences, North Carolina State University, Raleigh, NC 27695, USA*

Correspondence to Rodolphe Barrangou: *Department of Food, Processing and Nutrition Sciences, 400 Dan Allen Drive, Campus box 7624, North Carolina State University, Raleigh, NC 27695, USA.* *rbarran@ncsu.edu*
https://doi.org/10.1016/j.jmb.2018.09.013
Edited by Mazhar Adli

Abstract

CRISPR–Cas systems constitute the adaptive immune system of bacteria and archaea, as a sequence-specific nucleic acid targeting defense mechanism. The sequence-specific recognition and cleavage of Cas effector complexes has been harnessed to developed CRISPR-based technologies and drive the genome editing revolution underway, due to their efficacy, efficiency, and ease of implementation in a broad range of organisms. CRISPR-based technologies offer a wide variety of opportunities in genome remodeling and transcriptional regulation, opening new avenues for therapeutic and biotechnological applications. To repurpose CRISPR–Cas systems for these applications, the various elements of the system need to be first identified and functionally characterized in their native host. Bioinformatic tools are first used to identify putative CRISPR arrays and their associated genes, followed by a comprehensive characterization of the CRISPR–Cas system, encompassing predictions for guide and target sequences. Subsequently, interference assays and transcriptomic analyses should be performed to probe the functionality of the CRISPR–Cas system. Once an endogenous CRISPR–Cas system is characterized as functional, they can be readily repurposed by delivering an engineered synthetic CRISPR array or a small RNA guide for targeted gene manipulation. Alternatively, developing a plasmid-based system for heterologous expression of the necessary CRISPR components can enable exploitation in other organisms. Altogether, there is a wide diversity of native CRISPR–Cas systems in many bacteria and most archaea that await functional characterization and repurposing for genome editing applications in prokaryotes.

Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins represent the immune system of bacteria and archaea, and provide defense against invasive nucleic acids, DNA, or RNA from phages, plasmids, and other exogenous DNA elements [1,2]. CRISPR–Cas systems are widespread in bacteria and archaea, and the characterized CRISPR–Cas systems share common features, like the repeat-spacer arrays, that represent the vaccination card of the strain, and the associated signature Cas nucleases that perform cleavage in the invasive nucleic acid. Currently, two different classes, six types and 21 subtypes of CRISPR–Cas systems, have been described [3,4]. Type I and type II systems, with Cas3 and Cas9 as signature proteins, respectively, are the most predominant in nature [5,6] and have been heavily studied to understand the native components that enable these prokaryotic immune systems to target DNA. Elucidating the commonalities and differences between these two major types opens up the possibility to engineer and reprogram these CRISPR–Cas systems, leading to a wide range of applications from genome editing, genome remodeling, transcriptional regulation, and selectively targeting to eradicate pathogens [7]. Although the type II systems are less abundant in nature, their convenient simplicity with a unique Cas effector nuclease able to perform both DNA binding and cleavage makes this system much more portable and desirable for heterologous use in various organisms. Undoubtedly, Cas9 heterologous expression from the well-characterized and popular *Streptococcus pyogenes* type II system (SpyCas9) has become a powerful genetic tool to perform genome editing in a broad range of organisms

from bacteria [8] to human cells [9] and also to perform transcriptional regulation in bacteria and eukaryotes [10–14]. However, in bacteria and archaea, when a CRISPR–Cas system is present in the chromosome, the endogenous system can be conveniently reprogrammed for a wide range of applications, with fewer limitations than the heterologous expression of CRISPR–Cas9 plasmid-based delivery systems [8]. Indeed, delivering a small CRISPR array enables to reprogram the endogenous Cas machinery with simplicity and efficiency, with special relevance to recalcitrant strains where plasmid sizes drastically limit transformation efficiencies.

Nonetheless, before an endogenous CRISPR–Cas system can be repurposed for any application or even to develop a novel tool, the complete CRISPR–Cas locus of the strain of interest needs to be identified in the genome and fully functionally characterized. Here, we described how to (i) identify CRISPR–Cas systems with currently available bioinformatics tools, (ii) characterize their functionality based on transcriptomic data and interference assays, and (iii) repurpose endogenous and develop heterologous systems for a wide range of applications in bacteria.

CRISPR Biology and Mechanism of Action

CRISPR–Cas systems represent the adaptive immune system, which is DNA encoded and RNA mediated, providing protection against invasive nucleic acids (DNA and RNA) based on sequence-specific recognition [1,15]. CRISPR–Cas systems rely on the existence of a nuclease that is able to cleave DNA and a repeat-spacer CRISPR array that defines the target (s) of the nuclease; altogether these elements are typically co-located within an organized cluster in the host chromosome (DNA encoded) (Fig. 1). Distinctly, class I systems encompass a group of Cas proteins that constitute the Cascade (CRISPR-associated complex for antiviral defense) complex, whereas class II systems are conveniently reduced to a single multifunctional protein with the ability to bind and cleave DNA [3,4] (Fig. 1). The class I CRISPR–Cas systems, including type I, type III, and type IV, are the most abundant in nature, and among them, type I is the most predominant [5]. In contrast, class 2 encompasses type II, type V, and type VI, with type II the most explored to date [6]. While type III and type VI target RNA, the characterized types I, II, and V target DNA, with the type IV target still unclear, illustrating how both class I and class II include CRISPR–Cas systems that are able to target both DNA and RNA [3,4]. Type I and type II systems have been heavily studied and will be our focus in the subsequent sections.

As mentioned previously, type I systems are defined by a signature protein, Cas3, with helicase and exonuclease activity, which is able to nick and digest DNA in combination with the Cascade complex, whereas type II systems hinge on a single effector protein, Cas9, with endonuclease activity (Fig. 1). All CRISPR–Cas types and subtypes include Cas1 and Cas2 proteins, which are involved in the acquisition of new spacers from invasive elements [16,17]. Regarding the CRISPR arrays, the length of the repeat sequence varies from an average of 28 nt in type I and 36 nt in type II systems, with variability among CRISPR subtypes. On the other hand, the spacers are hypervariable regions and a defining feature of the CRISPR locus as they represent the immunization record of each strain [18]. Consequently, the spacers can be used for genotyping of strains, which is useful for the study of many pathogenic bacterial isolates [19–22]. Analyzing spacer sequences and matches to public databases reveals spacer origin and enables the determination of key targeting features such as the protospacer adjacent motif (PAM), a 3- to 5-nt conserved sequence that is recognized by the nuclease before cleaving the target nucleic acid [18]. The PAM is often located at the 5′-end of the protospacer in type I systems and at the 3′-end in type II systems. The PAM sequence recognized by the nuclease is dependent on the nuclease amino acid sequence and therefore varies between systems [23,24].

The CRISPR–Cas immune system mechanism of action encompasses three main steps for DNA-encoded, RNA-mediated, nucleic-acid targeting. During the first step, *acquisition*, Cas1 and Cas2 integrate new spacers by copying sequences from exogenous invasive nucleic acids and iteratively incorporating them as novel spacers at the 5′-end of the CRISPR array, close to the leader region (Fig. 1). The second step is the *expression* of the CRISPR array, which leads to the generation of the premature CRISPR RNA (pre-crRNA) that contains all the repeat-spacers in a single RNA molecule (RNA mediated). The pre-crRNA is then processed by Cas6 in type I systems (Cas5 in type I-C) or RNase III and Cas9 in type II systems, leading to the generation of the mature crRNA (CRISPR RNA) (Fig. 1). The last step is *interference* against the complementary invasive nucleic acid. In type I systems, the crRNA interact with the Cascade complex and generates the effector complex that recognizes and binds the invasive double-stranded DNA based on the spacer–protospacer match, recruiting Cas3 to nick and cleave target dsDNA. The base pairing between the crRNA and the complementary DNA target leads to the formation of an R loop on the non-target strand. Then, Cas3 exonuclease activity nicks the non-target strand within the protospacer and subsequently degrades this strand unidirectional in the 3′ to 5′ direction in an ion- and ATP-dependent manner [25–29]. In type II systems, the repeat sequence of the crRNA is complementary to the anti-repeat region of the trans-activating CRISPR RNA (tracrRNA) allowing the

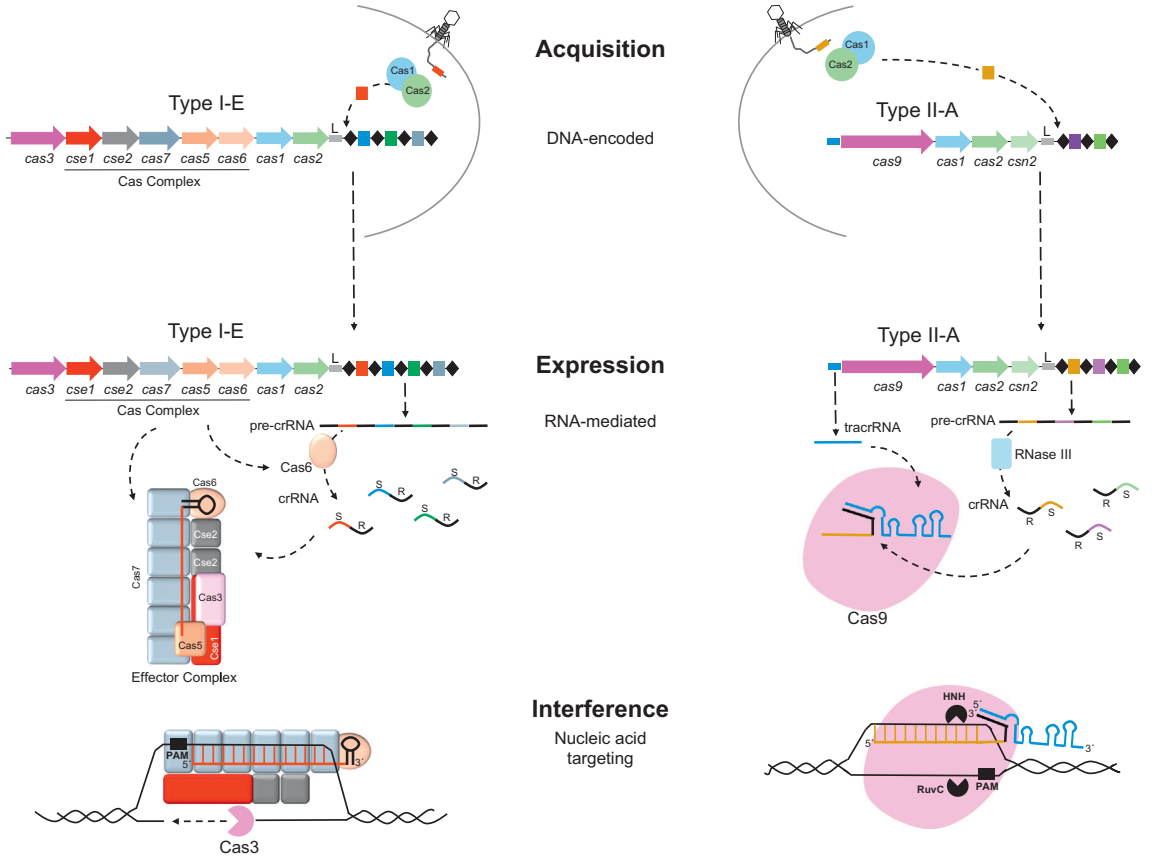

Fig. 1. CRISPR–Cas systems biology and mechanism of action. Class I CRISPR–Cas systems represented by type I-E (left) and class II represented by type II-A (right). CRISPR–Cas systems are DNA-encoded, RNA-mediated, and nucleic-acid targeting machines that work through a three step molecular mode of action. During the *acquisition* step, Cas1 and Cas2 copy and paste invasive DNA sequence that is added iteratively as a new spacer at the 5′-end of the CRISPR array, close to the leader region. During *expression*, the CRISPR array is transcribed as a pre-crRNA and then processed into a mature interfering crRNA by Cas6 in type I (Cas5 in type I-C) or RNaseIII and Cas9 in type II systems. Finally, the *interference* stage of nucleic acid targeting is mediated by the guide RNAs (crRNA in type I and duplex crRNA:tracrRNA in type II) that drives the Cas effector proteins toward sequences complementary to the spacer to selectively recognize and cleave DNA after PAM recognition.

formation of a duplex crRNA:tracrRNA (the dual-guide RNA) [30]. This dual-guide RNA interacts with Cas9 to guide the nuclease to the target, recognize double-stranded DNA, and perform a double-strand blunt cleavage, after PAM recognition, due to the action of two nuclease domains, HNH and RuvC, that each nicks one target strand precisely 3 nt from the 3′ edge of the target sequence [30,31] (Fig. 1).

Repurposing CRISPR–Cas Systems

In silico characterization of CRISPR–Cas elements

The CRISPR–Cas locus needs to first be identified in the genome and fully characterized prior to repurposing for any application. Initial steps include the identification of the *cas* genes and the repeat-spacer CRISPR array, usually encoded together in the same region of the genome as a cluster. Several algorithms have been developed to search for repeat regions that could represent a CRISPR array to facilitate the prediction and analysis of CRISPR loci. The main tools developed to date are command-line executable programs like CRISPR recognition tool (CRT), CRISPRdisco, CRISPRviz, MinCED, CRASS, PILER-CR [32–36], Web applications like CRISPRFinder, CRISPI, and more recently CRISPRCasFinder and CRISPRDetect [37–40]. Meanwhile, these various softwares are based on the identification of short palindromic repeat sequences, separated by similar distance (regularly interspaced), that are clustered in the same region of the genome, representing a putative CRISPR array. The differential performance between these softwares is variable with regard to accuracy, precision, speed, recall and quality [33]. These differences are primarily

based on the mismatches allowed between repeats and *e*-value thresholds (scoring matrices), although they all manage assembled genomes but have variable and limited potential for metagenomic data. Notably, the CRT algorithm has been one of the most widely used and has been implemented as a plugin ready to use in several genome browsers to make it accessible to non-command line experts. Once the CRISPR array is detected, the *cas* genes can be located in their vicinity (Fig. 2), although there can be several arrays located in trans in the genome. The plethora of draft genomes that are poorly assembled pose a challenge to CRISPR and *cas* co-location and characterization, which is compounded by the fact that CRISPR loci, as repeat-containing elements, are difficult to assemble in the first place, and often can be found at the edge of contigs. The Cas proteins can be identified through BLAST algorithms [41] using other previously identified Cas proteins as a query, although sequence conservation varies across CRISPR–Cas types and subtypes. Recently, a command line pipeline, CRISPRdisco, was developed to identify CRISPR–Cas systems, encompassing CRISPR arrays and Cas proteins, in genomic data sets [34]. When the signature *cas* genes and the associated Cas proteins are annotated adjacent to the repeat-spacer array, the CRISPR locus will be displayed as a single feature as represented in Fig. 2. The CRISPR arrays are depicted with diamonds that represent the repeats and with squares representing the unique spacers. A BLAST analysis of the spacer sequences against the databases may elucidate the origin of the invasive nucleic acid (phage or plasmids) allowing users to potentially predict the PAM [18,42]. To avoid a manual BLAST of the spacers, webserver tools like CRISPRtarget were developed [43]. As aforementioned, the PAM is located at the 5′-end of the protospacer in type I systems and at the 3′-end in type II systems (Fig. 2). Therefore, the 8- to 10-nt sequences flanking the protospacer are manually extracted and used to identify and predict the conserved PAM based on a frequency chart using WebLogo [44], which represents the conservation of each nucleotide at each position (Fig. 2). Moreover, two additional tools, PAM-SCANR and PAM wheel, were developed to elucidate functional PAMs and visualize PAM sequences and their relative interference potentials [45].

In type I systems, the crRNA represents the guide RNA that interacts with the Cas complex to target the complementary sequence (Fig. 2). The hairpin structure encoded within the repeat due to the palindromic nature of the CRISPR sequence can be predicted *in silico* using NUPACK [46]. Type II systems also require the prediction of the tracrRNA as another essential component of the CRISPR–Cas systems. The tracrRNA has a region complementary to the repeat sequence of the crRNA that allow to create the duplex crRNA:tracrRNA. Accordingly, the repeat sequence can be used to search for partial complementarity to the CRISPR locus to identify the tracrRNA, usually located in close proximity to the *cas9* gene or the CRISPR array [23,24]. A detailed protocol to predict the tracrRNA, including step by step process and troubleshooting was previously described by Briner and co-workers [47]. Once the tracrRNA sequence is identified, the interaction between the repeat portion of the crRNA and the tracrRNA anti-repeat can be predicted using NUPACK tool [46] with the set up for two strands, a maximum complex size of two and 1 μM of each strand [47]. The resulting structure will display the predicted interaction of the dual-guide crRNA:tracrRNA and the hairpin structures contained within the tracrRNA (Fig. 2).

Evaluation of CRISPR–Cas activity and functionality

The characterization of each component of the CRISPR locus of interest is essential for repurposing CRISPR–Cas systems for further applications. The transcription of CRISPR–Cas elements and the activity of the various promoter regions that guide the expression of the CRISPR locus can be inferred from RNA-seq transcriptome data. Expression of the CRISPR–Cas locus is growth condition dependent; nonetheless, these loci are typically transcribed during planktonic growth and have been repeatedly detected in RNA-seq expression studies, suggesting that the system is potentially active and functional in standard growth conditions. A monocistronic RNA profile is usually detected for *cas3*, whereas there is a typical polycistronic RNA for the Cascade complex, suggesting the existence of two different promoters for this locus (Fig. 3a). The mRNAseq data typically provide the expression profile of the CRISPR array, although small RNA-seq analysis is desired to elucidate the expression of each spacer-repeat pair, and determine the boundaries of processed, mature crRNAs. Usually, the most highly transcribed spacer is the most recently acquired, which is located in the 5′-end of the CRISPR array close to the leader sequence (Fig. 3a). Indeed, the first few spacers, representing recent immunization events, are typically more transcribed than ancestral spacers that provide resistance against more historically but less currently relevant foes. In type II systems, the RNA-seq data will also help to determine the tracrRNA location, sequence, and expression profile, as it is transcribed independently from the other CRISPR–Cas components (Fig. 3b). Noteworthy, determining the 3′ composition of the tracrRNA, especially the terminal hairpins, is not trivial to predict informatically, and small RNA characterization is instrumental in correctly defining the boundaries of tracrRNA molecules.

Finally, once the *in silico* characterization and predictions are complete, the functionality of the CRISPR systems and the correct identification or

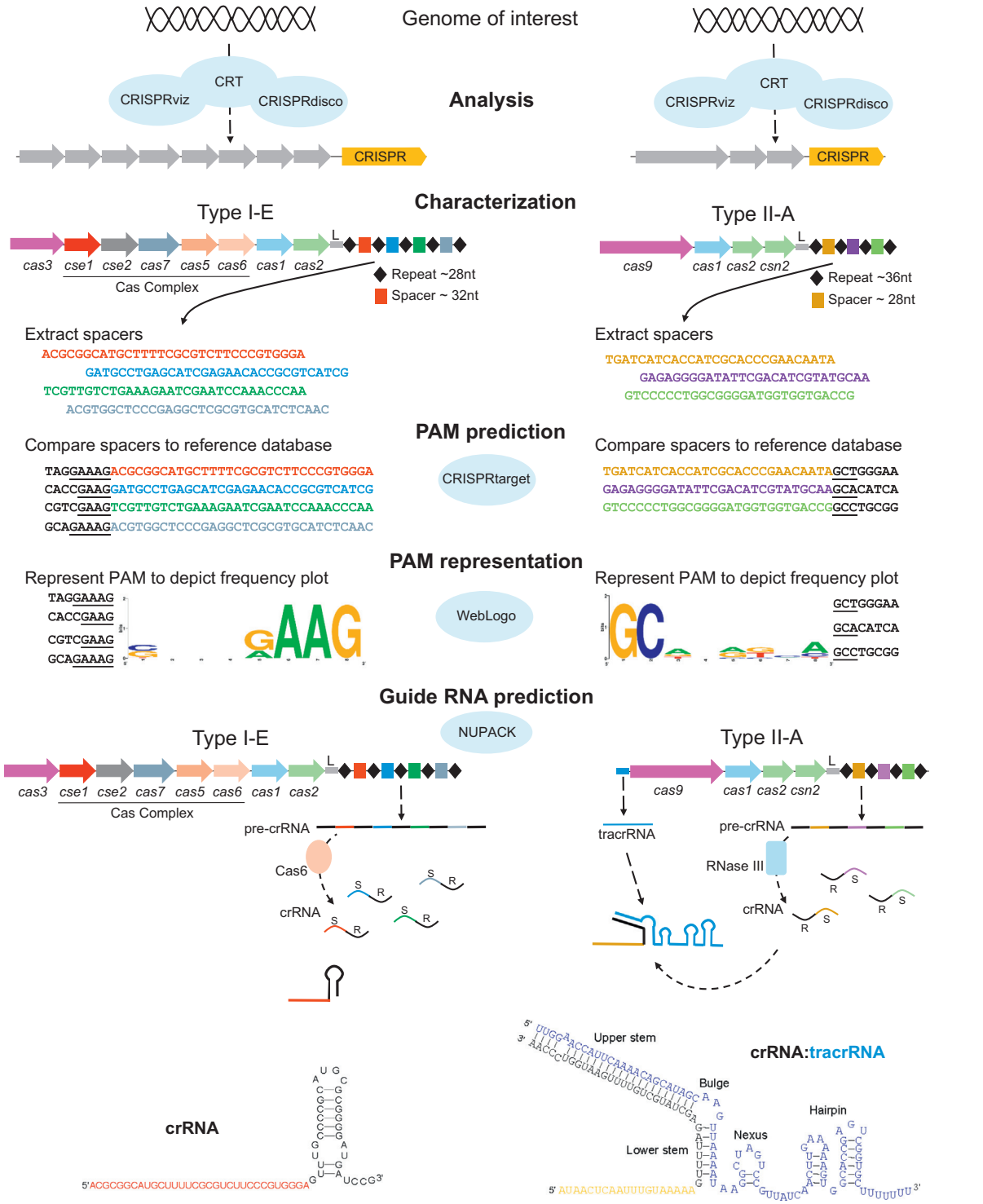

Fig. 2. Identification and characterization of CRISPR–Cas systems. Workflow of the steps to identify and characterize new CRISPR–Cas systems in genomic DNA from bacteria and archaea using *in silico* tools. Initial steps of identification, prediction and characterization of potential CRISPR–Cas systems based on the presence of CRISPR arrays and *cas* genes (left: type I, right: type II). The spacers are used for PAM prediction based on spacer–protospacer identification in databases, followed by the representation of the conserved sequence flanking the protospacer. Prediction of the hairpin structures of the crRNA in type I systems or the interaction between the crRNA and tracrRNA (dual-guide crRNA:tracrRNA) in type II systems is shown on the bottom panel. Blue light circles represent the *in silico* tools that can be used in each step.

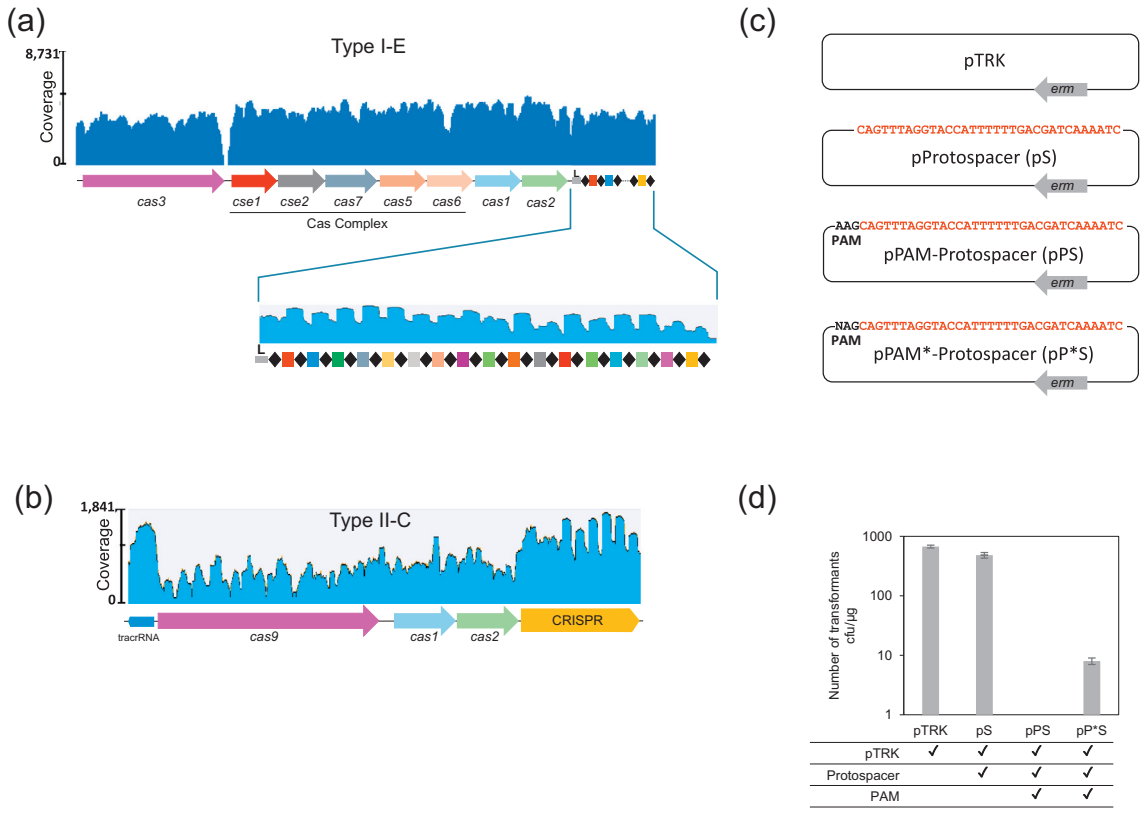

Fig. 3. CRISPR activity and functionality. The CRISPR locus activity can be predicted based on transcriptome analysis encompassing mRNA-seq to probe *cas* genes expression and small RNA-seq to determine the transcription level and sequence composition of mature crRNA from CRISPR arrays (a). In type II systems, the RNA-seq data will also display the profile and composition of the tracrRNA (b). The functionality of the CRISPR locus is probed by plasmid interference assays, targeting protospacer and PAM sequences (e.g., type I system, PAM: 5′-AAG-3′) cloned into a target plasmid, including a mutated PAM* (5´-NAG-3′) (c) and checking survival of clones under selective antibiotic pressure (erythromycin) (d).

validation of the PAM should be confirmed by experimental data. The CRISPR functionality and PAM validation analysis can be performed using phage infection or plasmid interference assays [48]. The invasive DNA should carry the predicted PAM in the vicinity of a protospacer that matches a spacer contained within the CRISPR locus. The most highly expressed spacers are typically the most desirable targets to clone along with the predicted PAM in a plasmid to probe the functionality of the CRISPR system by plasmid interference assays (Fig. 3c). This enables the assessment of the functionality of the endogenous CRISPR array and *cas* genes in their native context (including promoters and regulatory elements as well as ancillary proteins). After the transformation of the target plasmid, an active CRISPR–Cas system will recognize the invasive DNA *via* sequence-based homology between the CRISPR spacer and the protospacer on the target plasmid, and once the correct PAM is recognized and the adjacent DNA sequence interrogated, the nuclease will cleave the plasmid. As a consequence of plasmid targeting, no survivor will be obtained in selective medium containing the antibiotic that corresponds to the antibiotic marker encoded on the plasmid (Fig. 3d). A plasmid with the protospacer but without the PAM should be transformed in parallel as a control, and colonies will be obtained in the selective medium as well as for the cells transformed with the empty vector. These results will demonstrate the functionality of the endogenous CRISPR–Cas system and validate the predicted PAM. Also, altered PAM can be cloned with the same protospacer for further characterization of the PAM or selection of the most efficient PAM based on targeting efficiencies, when multiple PAMs can be recognized by the same Cas effector nucleases (Fig. 3d). Once the functionality of the CRISPR system is confirmed, the endogenous system can be conveniently reprogrammed for a wide variety of applications as discussed below.

Harnessing endogenous CRISPR–Cas systems

The CRISPR–Cas systems of bacteria and archaea can be repurposed to target the host genome, enabling a wide range of applications. The ability to introduce a

small CRISPR array to guide the nuclease to the desired chromosomal target opens the possibility to perform genome remodeling in an accurate, reliable, and efficient manner. For this purpose, a small CRISPR array should be cloned on a plasmid under the expression of a promoter, usually the native leader of the CRISPR array (Fig. 4a). This cloning strategy can be performed in one step using a synthesized gene block that contains the four components (Leader-Repeat-Spacer-Repeat) (Fig. 4a.1). However, if several individual targets are going to be in play, it is recommended to construct an initial plasmid cloning a gene block that contains the leader sequence and two repeats, without spacer, and two restriction sites in opposite direction between the repeats to offer *plug and play* and flexible reprogramming of user-selected spacers (Fig. 4a.2). The pcrRNA plasmid can then be digested, and the desired targeting spacer will be incorporate between the repeats using annealing oligonucleotides carrying overhangs compatible with the digested plasmid (Fig. 4a.2). The spacer that defines the target should be a 20- to 30-nt sequence located downstream of the 3′ edge of the PAM for type I or upstream of the 5′ edge of the PAM for type II systems. Once the spacer is defined, the absence of homology, in particularly within the seed sequence of the spacer, with other regions of the chromosome should be checked to reduce the possibility of off targets effects.

The transformation of the plasmid inside the cell leads to the expression of the CRISPR array cloned in the plasmid and guides the Cas machinery toward the desired target, binds the complementary dsDNA, and performs a blunt double-stranded cleavage in type II systems or single strand cleavage in type I systems [27] (Fig. 1). The DNA repair mechanisms in prokaryotes are relatively poor and limited compare to their eukaryotic counterparts, and typically not efficient enough to repair the DNA damage caused by the nuclease; consequently, targeting bacterial genomes is highly lethal and often leads to bacterial cell death (Fig. 4b). Bacterial DNA repair is primarily based on

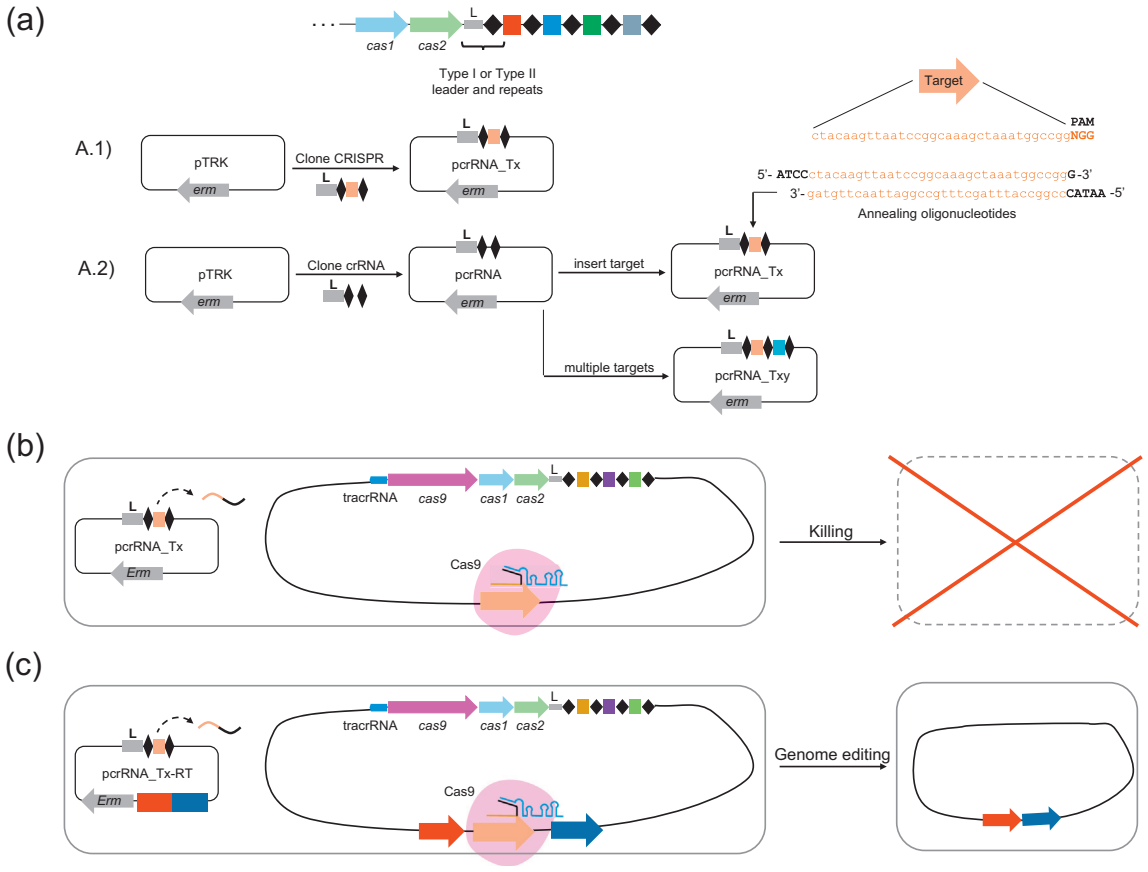

Fig. 4. Cloning strategy to repurpose endogenous CRISPR–Cas systems. Plasmid-based delivery (erythromycin resistance marker) of a small CRISPR array under the expression of the leader allows the repurposing of the endogenous CRISPR–Cas system (a). Once the plasmid is transformed, the Cas machinery is guided to target the desired sequence, enabling lethal self-targeting that leads to killing if DNA repair mechanisms are not efficient (b). A repair template can be provided in the targeting plasmid as donor DNA to allow the DNA repair machinery to overcome CRISPR cleavage, introducing the desired mutation designed within the repair template (c).

homologous recombination (HR) with RecA as the main recombinase, which favors DNA repair using a homologous template to restore the genetic locus impacted at the cleavage site. Therefore, incorporating a repair template within the targeting plasmid will provide an opportunity for DNA repair mechanisms to overcome the damage *via* HR, while introducing engineered mutation, providing programmable genome editing (Fig. 4c).

CRISPR Applications in Bacteria

Endogenous systems

The ability to reprogram endogenous CRISPR–Cas systems in bacteria and archaea leads to a wide range of applications and has several advantages compared to heterologous expression. Repurposing the endogenous Cas machinery to deliver a crRNA targeting the chromosome will lead to self-targeting and, without providing a repair template, selective killing. This application can be used as a screening method to select rare variants among a population present in nature, killing the wild-type genotype and enabling the selection of natural rare variants that have mutations in the targeted sequence, without the generation of genetically modified organisms (Fig. 5a). Another option to obtain genetic variants is targeting dispensable genomic islands to perform genome remodeling, by selectively enabling the survival of rare variants that have deleted the target region, which can occur by recombination between transposons, and select for the loss of sometimes large genomic islands (over 100 kb) devoid of essential genes [49,50]. Self-targeting by CRISPR can also be combined with conventional HR-directed gene replacement systems when providing the corresponding engineered donor DNA that will serve as a repair template or with single-stranded DNA oligonucleotide recombineering for direct selection of recombinants carrying the desired mutation [51,52], while killing the population which retained the wild-type genotype.

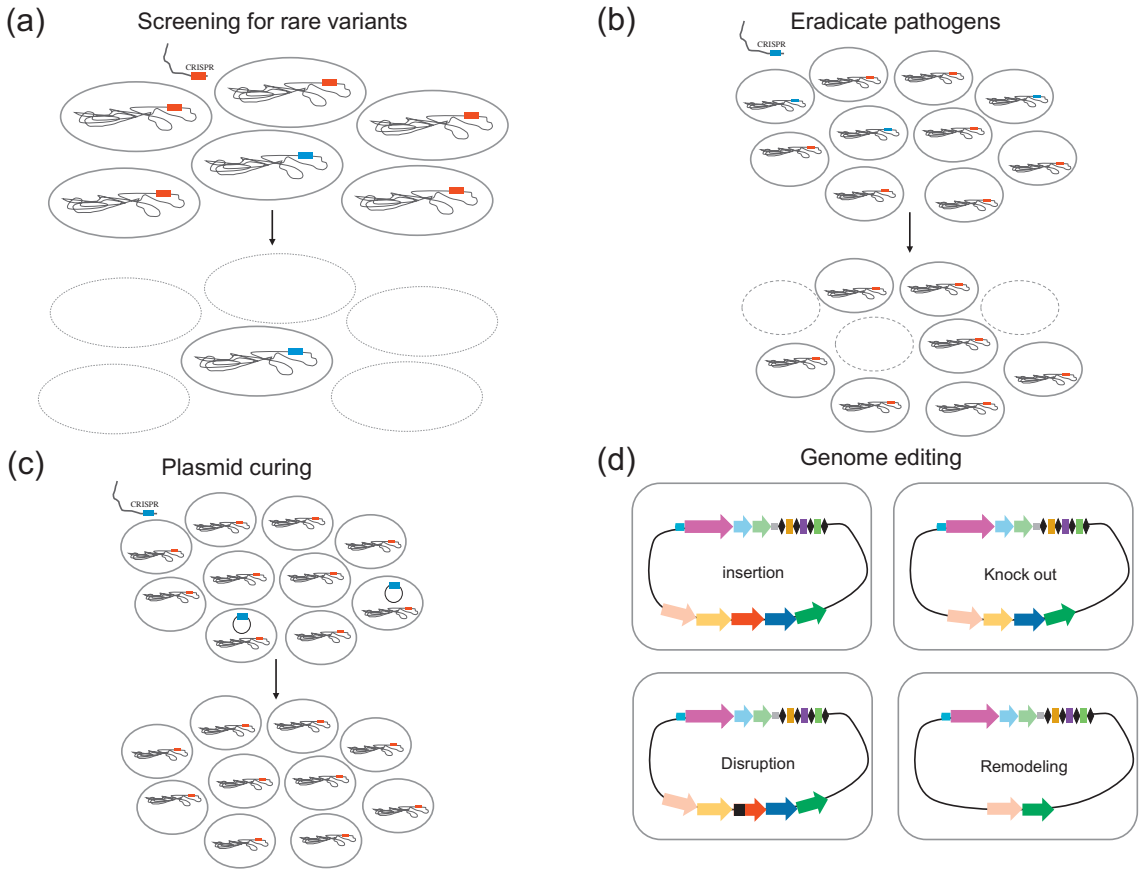

Fig. 5. CRISPR applications. Delivering a crRNA to repurpose the endogenous CRISPR–Cas systems allows to screen for natural rare variants targeting the wild-type genotype (a) and also to selectively eradicate pathogens in a mixed bacterial population as a sequence-specific antimicrobial strategy (b) or to perform plasmid curing to eliminate virulence genes or antibiotic resistance genes usually encoded on plasmids (c). Providing a repair template allows to perform different genetic engineering strategies such as gene insertion, knockout, gene disruption, or large deletions (d).

CRISPR targeting of specific bacterial strains can be applied to selectively eradicate pathogens from a mixed population using either type I [53] or type II system [54,55] for developing novel sequence-specific antimicrobials and to alter the bacterial population within a defined microbiome (Fig. 5b). In addition, plasmid curing can be performed by targeting the antibiotic resistance genes (or virulence genes) present in pathogenic bacteria that are encoded on plasmids (Fig. 5c). Eliminating antibiotic resistance genes leading to developing-sensitive strains will help reduce the occurrence and dissemination of antibiotic resistance genes in various habitats and environmental niches. Moreover, the capability to perform selective targeting in a mixed population will allow to reshape the human microbiome, modulating the occurrence and diversity of certain species toward a healthier microbiota. Plasmid-based delivery of crRNA is a powerful bench tool as a proof of concept that can even work in some *in vitro* models, but needs to be improved by alternative delivery methods to be applied *in vivo*. In this regard, phages have been used to deliver crRNA and repurpose endogenous systems to perform selectively killing of *Staphylococcus aureus* in murine models [54].

Finally, when a DNA template is provided to repair the DNA damage from CRISPR targeting, different genome editing strategies can be achieved, including gene insertion, knockout, gene disruption, or genome remodeling with large-scale deletions (Fig. 5d). To date, relatively few studies have been performed to repurpose endogenous CRISPR–Cas systems for genome editing in bacteria or archaea. For instance, the native type I-A and type III-B systems have been redirected in the archaeon *Sulfolobus islandicus* for gene deletion, insertion and point mutations, combining the target and the donor DNA template in one plasmid [56]. Similarly, the endogenous type I-B CRISPR–Cas system was repurposed for precise genome editing in *Haloarcula hispanica* for gene deletion, tagging, and single-nucleotide substitution, also using a unique plasmid carrying both the CRISPR target and the repair template, even with two different loci edited in one step [57]. Regarding transcriptional control, the endogenous type I system has been used for gene repression in archaea [58] and in *Escherichia coli,* first by eliminating the Cas3 nuclease activity [59,60]. In this case, when the crRNA is delivered to the cells, the Cas complex is guided to irreversibly bind the desired target, without cleavage due to the absence of Cas3, resulting in blocking the progress of RNA polymerase and repressing transcription. The exploitation of the endogenous CRISPR systems for genome editing in bacteria has been probed in clostridia, using the native system type I-B in *Clostridium pasteurianum* [61] and in *Clostridium tyrobutyricum* [62]. Moreover, Pyne and co-workers [61] showed that the use of endogenous system I-B was more efficient to perform genome editing than the widely used *S. pyogenes* Cas9 in heterologous expression. Zhang and co-workers [62] reported that type I-B can be used for gene deletion or gene integration with 100% efficiency. A single plasmid was also used in this study to co-deliver the CRISPR target and the repair template that provides the donor DNA, with two deletions performed in a single step [62]. Interestingly, aside from *S. pyogenes,* there has not been other report to date utilizing native type II CRISPR–Cas machinery for editing native bacterial genomes. This is presumably due to the lack of non-homologous end-joining (NHEJ) pathway in most prokaryotes for repair of double-stranded DNA break generated by Cas9, although inactivation of one of the two nuclease domains in the native Cas9 will very likely empower genome engineering applications with the otherwise intractable endogenous type II systems. Meanwhile, Martel and Moineau [63] exploited the endogenous type II-A system in *Streptococcus thermophilus* for generating gene deletion, insertion, and point mutation in streptococcal lytic phage, demonstrating the potential of CRISPR–Cas system as a powerful tool for virulent phage engineering and the study of phage-host interaction.

Heterologous system

CRISPR–Cas technology is not limited to the strains harboring endogenous systems. Indeed, when CRISPR is not naturally present in a given strain, CRISPR–Cas systems can be introduced on a plasmid with the required components to perform genome editing. In the case of type I systems, the Cascade complex, the Cas3 nuclease, and the crRNA required for delivery pose a limitation in the cloning process due to the size of the various CRISPR–Cas system components and the total cargo to package in a single plasmid. Kiro and co-workers [64] used a two-plasmid-based systems to clone independently the Cascade complex and Cas3, reducing the size of the final constructions, although creating the challenge of requiring two distinct selective markers.

In the case of type II systems, Cas9 nuclease together with the native dual-guide crRNA and tracrRNA (crRNA:tracrRNA), or the engineered single-guide RNA (sgRNA) have been developed as a plasmid-based delivery technology to perform genome editing in a wide variety of organisms from bacteria to human cells [7]. The absence of a Cascade complex in type II systems makes this strategy more portable than using heterologous type I systems. In this regard, the portable plasmid-based Cas9 has been used to perform genome editing in bacteria including *E. coli*, *Clostridium*, *Bacillus subtilis* [65–70], as well as gene editing in virulent phages of *Lactococcus lactis* and *Klebsiella pneunoniae* [71,72]. Moreover, different versions of Cas9 have been developed being able to nick only one DNA strand (nickase variant or nCas9) by inactivating of either the RuvC or HNH nickase domain, thereby reducing the lethality of cleavage and

increasing the genome editing efficiency, and also a deactivated version (dCas9), with both RuvC and HNH nickase domains mutated, that binds but does not cleave DNA, providing a DNA binding protein able to block RNA polymerase, thus acting as a transcriptional regulator [10,12,14,73–75].

In the heterologous plasmid-based delivery, the nuclease and the sgRNA should be under the control of separate promoters that regulate their expression. The use of host-derived inducible promoters, weak or strong promoters as well as the copy number of the plasmid will have an impact on the overall efficiency of the heterologous tool. To perform genome editing in bacteria, the sgRNA is generally expressed under a strong promoter, and if possible in a high copy plasmid, which is especially crucial for providing sufficient copies of the repair template that serves as the donor for DNA repair. On the other hand, in cases where the native Cas9 is employed, the nuclease should be placed under the regulation of a low-expression promoter or in a low-copy plasmid to minimize the toxicity of the Cas nuclease activity and therefore the lethality of the DNA damage, allowing the user to recover survivors that carry the desired mutation at higher frequencies. Alternately, the use of inducible promoters that regulate the nuclease expression is strongly desired to induce nuclease activity during optimal conditions to enhance gene editing efficiencies. Once the plasmid has been transformed into the strain of interest, which is then grown on the selective marker, the promoter of *cas9* can be switched on by adding the corresponding inducer (generally a substrate of sugar metabolism), thereby overcoming the challenges of transformation-induced stress and Cas nuclease-induced lethality. To date, heterologous expression of SpyCas9 has been performed in different ways, the latest with Cas9 or nCas9 and the sgRNA under different promoters, along with the repair template, all on a single plasmid. Several studies have also reported robust gene editing efficacy when adopting the SpyCas9 nickase variant, thus circumventing the limited indigenous DNA repair capacity [76–80]. The presence of NHEJ repair enzymes in some species of *Mycobacterium*, *Bacillus*, and *Streptomyces*, namely, the Ku protein and DNA ligase IV, has also been exploited to complement the CRISPR–Cas9 systems for HR-independent genome editing. For example, CRISPR–Cas9 systems have been employed efficiently for gene deletions in *Streptomyces coelicolor*, with or without editing template provided, relying on the native NHEJ pathway for repair of DNA cleavage [81]. The NHEJ pathway from *Mycobacterium* has also recently been heterologously expressed in *E. coli* to complement the CRISPR–Cas9 editing system, providing an efficient strategy for deletion of large chromosomal fragment without the requirement of a donor repair template [82,83]. In this regard, constitutively high expression of the NHEJ pathway on a high copy plasmid is imperative for efficient repair of Cas9 cleavage, as chromosomal expression of the pathway, either native or in heterologous hosts, was not sufficient to counteract the lethality of the Cas9 endonuclease activity [80,82,84]. It is noteworthy that the NHEJ repair is evidently inhibited by the Cas protein Csn2 of type II-A system, thus limiting the heterologous application of this efficient DNA repair pathway for genome engineering in prokaryotes harboring type II-A system [85].

The generation of portable and efficient CRISPR–Cas tools has opened new avenues for applications in potentially any organism of interest. However, the application in bacteria is still limited by transformation efficiencies, creating a need for compatible plasmid vectors and novel delivery technologies to broaden the application to a wide range of species. Delivery methods need to be improved as to limit damage (i.e., electroporation) and increase throughput (i.e., efficiency, load capacity) while providing efficient delivery of DNA molecules. Electroporation and heat shock are the most common approaches in bacteria, whereas adenovirus, nanoparticles, and nanovesicles have been successfully used in eukaryotes to deliver DNA. As previously mentioned, phage engineering is an alternative for delivery of CRISPR arrays or guides and affords the option to repurpose endogenous systems. However, delivering relatively large constructs like complete CRISPR–Cas systems can be challenging, especially in the case of class I systems. Furthermore, a major limitation for phage-based delivery is the often narrow host range and difficulties inherent to engineering bacteriophage genomes. Ultimately, genome editing outcomes hinge on the fundamental understanding of prokaryotic DNA repair mechanisms. Overcoming these limitations becomes even more important when delivering heterologous CRISPR–Cas systems as the plasmid sizes average 10–11 kb, due the size of Cas9 (~4.5 kb), not to mention the occasional problematic Cas9 instability during the cloning process. Therefore, when possible, repurposing the endogenous systems in strains possessing native CRISPR–Cas systems is arguably a desirable and promising alternative to delivering heterologous CRISPR–Cas.

Conclusions

CRISPR–Cas systems offer new avenues for genome editing in a wide range of organisms, from bacteria to eukaryotic cells. Overall, CRISPR technologies are primed to increase the ease and efficiency of genetic alterations, reducing time and technical difficulties to modify recalcitrant organisms. Although CRISPR-based technologies have been rapidly and broadly applied in eukaryotes, relatively few applications have been documented in bacteria. This is somewhat perplexing given the broad occurrence of

diverse CRISPR–Cas systems in many bacteria and most archaea, especially given the relative ease with which both heterologous and endogenous systems can be harnessed for genome remodeling, for transcription regulation, and also to selectively eradicate pathogens. Repurposing endogenous systems in particular opens new possibilities for the genesis of novel bacterial strains of medical, agricultural, and biotechnological potential and will quickly expand the CRISPR toolbox for next-generation genome editing and beyond.

Acknowledgments

The authors would like to thank their colleagues and support from NC State University and the North Carolina Agricultural Foundation.

Conflicts of Interest: R.B. is a co-inventor on several patents regarding CRISPR–Cas systems and their uses. R.B. is also a co-founder and shareholder of Intellia Therapeutics and Locus Biosciences, and a shareholder of DuPont and Inari. Claudio Hidalgo Cantabrana co-founder and SAB member of Microviable Therapeutics.

Received 30 July 2018;
Received in revised form 19 September 2018;
Accepted 19 September 2018
Available online 24 September 2018

Keywords:
CRISPR–Cas systems;
CRISPR characterization;
endogenous systems;
genome editing

Abbreviations used:
CRISPR, clustered regularly interspaced short palindromic repeats; Cascade, CRISPR-associated complex for antiviral defense; PAM, protospacer adjacent motif; pre-crRNA, premature CRISPR RNA; tracrRNA, trans-activating crRNA; HR, homologous recombination; NHEJ, non-homologous end-joining; sgRNA, single-guide RNA.

References

[1] R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, et al., CRISPR provides acquired resistance against viruses in prokaryotes, Science 315 (2007) 1709–1712.
[2] R. Barrangou, L.A. Marraffini, CRISPR–Cas systems: prokaryotes upgrade to adaptive immunity, Mol. Cell 54 (2014) 234–244.
[3] K.S. Makarova, F. Zhang, E.V. Koonin, SnapShot: class 1 CRISPR–Cas systems, Cell 168 (2017) 946-e1.
[4] K.S. Makarova, F. Zhang, E.V. Koonin, SnapShot: class 2 CRISPR–Cas systems, Cell 168 (2017) 328-e1.
[5] K.S. Makarova, Y.I. Wolf, O.S. Alkhnbashi, F. Costa, S.A. Shah, S.J. Saunders, et al., An updated evolutionary classification of CRISPR–Cas systems, Nat. Rev. Microbiol. 13 (2015) 722–736.
[6] S. Shmakov, A. Smargon, D. Scott, D. Cox, N. Pyzocha, W. Yan, et al., Diversity and evolution of class 2 CRISPR–Cas systems, Nat. Rev. Microbiol. 15 (2017) 169–182.
[7] R. Barrangou, J.A. Doudna, Applications of CRISPR technologies in research and beyond, Nat. Biotechnol. 34 (2016) 933–941.
[8] C. Hidalgo-Cantabrana, S. O'Flaherty, R. Barrangou, CRISPR-based engineering of next-generation lactic acid bacteria, Curr. Opin. Microbiol. 37 (2017) 79–87.
[9] A.C. Komor, A.H. Badran, D.R. Liu, CRISPR-based technologies for the manipulation of eukaryotic genomes, Cell 169 (2017) 559.
[10] X. Dong, Y. Jin, D. Ming, B. Li, H. Dong, L. Wang, et al., CRISPR/dCas9-mediated inhibition of gene expression in *Staphylococcus aureus*, J. Microbiol. Methods 139 (2017) 79–86.
[11] L.A. Gilbert, M.H. Larson, L. Morsut, Z. Liu, G.A. Brar, S.E. Torres, et al., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes, Cell 154 (2013) 442–451.
[12] J.M. Peters, A. Colavin, H. Shi, T.L. Czarny, M.H. Larson, S. Wong, et al., A comprehensive, CRISPR-based functional analysis of essential genes in bacteria, Cell 165 (2016) 1493–1506.
[13] L.S. Qi, M.H. Larson, L.A. Gilbert, J.A. Doudna, J.S. Weissman, A.P. Arkin, et al., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, Cell 152 (2013) 1173–1183.
[14] C. Zhao, X. Shu, B. Sun, Construction of a Gene Knockdown System Based on Catalytically Inactive ("Dead") Cas9 (dCas9) in *Staphylococcus aureus*, Appl. Environ. Microbiol. 83 (2017) pii: e00291-17.
[15] V.J. Price, W. Huo, A. Sharifi, K.L. Palmer, CRISPR–Cas and restriction-modification act additively against conjugative antibiotic resistance plasmid transfer in *Enterococcus faecalis*, mSphere 1 (2016).
[16] J.K. Nunez, P.J. Kranzusch, J. Noeske, A.V. Wright, C.W. Davies, J.A. Doudna, Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity, Nat. Struct. Mol. Biol. 21 (2014) 528–534.
[17] I. Yosef, M.G. Goren, U. Qimron, Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*, Nucleic Acids Res. 40 (2012) 5569–5576.
[18] F.J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system, Microbiology 155 (2009) 733–740.
[19] F. Almeida, M.I. Medeiros, D.D. Rodrigues, M.W. Allard, J.P. Falcao, Molecular characterization of *Salmonella typhimurium* isolated in Brazil by CRISPR-MVLST, J. Microbiol. Methods 133 (2017) 55–61.
[20] J.M. Andersen, M. Shoup, C. Robinson, R. Britton, K.E. Olsen, R. Barrangou, CRISPR diversity and microevolution in *Clostridium difficile*, Genome Biol. Evol. 8 (2016) 2841–2855.
[21] P.J. Freidlin, I. Nissan, A. Luria, D. Goldblatt, L. Schaffer, H. Kaidar-Shwartz, et al., Structure and variation of CRISPR and CRISPR-flanking regions in deleted-direct repeat region *Mycobacterium tuberculosis* complex strains, BMC Genomics 18 (2017) 168.

[22] S.M. Kovanen, R.I. Kivisto, M. Rossi, M.L. Hanninen, A combination of MLST and CRISPR typing reveals dominant *Campylobacter jejuni* types in organically farmed laying hens, J. Appl. Microbiol. 117 (2014) 249–257.
[23] K. Chylinski, A. Le Rhun, E. Charpentier, The tracrRNA and Cas9 families of type II CRISPR–Cas immunity systems, RNA Biol. 10 (2013) 726–737.
[24] A.E. Briner, R. Barrangou, Guide RNAs: a glimpse at the sequences that drive CRISPR–Cas systems, Cold Spring Harb Protoc 2016 (2016) (pdb top090902).
[25] Y. Huo, K.H. Nam, F. Ding, H. Lee, L. Wu, Y. Xiao, et al., Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation, Nat. Struct. Mol. Biol. 21 (2014) 771–777.
[26] R.N. Jackson, S.M. Golden, P.B. van Erp, J. Carter, E.R. Westra, S.J. Brouns, et al., Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*, Science 345 (2014) 1473–1479.
[27] T. Sinkunas, G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath, V. Siksnys, Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system, EMBO J. 30 (2011) 1335–1342.
[28] P.B.G. van Erp, A. Patterson, R. Kant, L. Berry, S.M. Golden, B.L. Forsman, et al., Conformational dynamics of DNA binding and Cas3 recruitment by the CRISPR RNA-guided Cascade complex, ACS Chem. Biol. 13 (2018) 481–490.
[29] B. Wiedenheft, E. van Duijn, J.B. Bultema, S.P. Waghmare, K. Zhou, A. Barendregt, et al., RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 10092–10097.
[30] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, Science 337 (2012) 816–821.
[31] M. Jinek, F. Jiang, D.W. Taylor, S.H. Sternberg, E. Kaya, E. Ma, et al., Structures of Cas9 endonucleases reveal RNA-mediated conformational activation, Science (2014) 343 (1247997).
[32] Minced: Mining CRISPRs in Environmental Datasets. https://sourceforge.net/projects/minced/.
[33] C. Bland, T.L. Ramsey, F. Sabree, M. Lowe, K. Brown, N.C. Kyrpides, et al., CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats, BMC Bioinf. 8 (2007) 209.
[34] A.B. Crawley, J.R. Henriksen, R. Barrangou, CRISPRdisco: An automated pipeline for the discovery and analysis of CRISPR-Cas systems, CRISPR J. 1 (2018) 171–181.
[35] R.C. Edgar, PILER-CR: fast and accurate identification of CRISPR repeats, BMC Bioinf. 8 (2007) 18.
[36] C.T. Skennerton, M. Imelfort, G.W. Tyson, Crass: identification and reconstruction of CRISPR from unassembled metagenomic data, Nucleic Acids Res. 41 (2013), e105.
[37] A. Biswas, R.H. Staals, S.E. Morales, P.C. Fineran, C.M. Brown, CRISPRDetect: a flexible algorithm to define CRISPR arrays, BMC Genomics 17 (2016) 356.
[38] D. Couvin, A. Bernheim, C. Toffano-Nioche, M. Touchon, J. Michalik, B. Neron, et al., CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins, Nucleic Acids Res. 46 (2018) W246–W251, https://doi.org/10.1093/nar/gky425.
[39] I. Grissa, G. Vergnaud, C. Pourcel, CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats, Nucleic Acids Res. 35 (2007) W52–W57.
[40] C. Rousseau, M. Gonnet, M. Le Romancer, J. Nicolas, CRISPI: a CRISPR interactive database, Bioinformatics 25 (2009) 3317–3318.
[41] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[42] H. Deveau, R. Barrangou, J.E. Garneau, J. Labonte, C. Fremaux, P. Boyaval, et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*, J. Bacteriol. 190 (2008) 1390–1400.
[43] A. Biswas, J.N. Gagnon, S.J. Brouns, P.C. Fineran, C.M. Brown, CRISPRTarget: bioinformatic prediction and analysis of crRNA targets, RNA Biol. 10 (2013) 817–827.
[44] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.
[45] R.T. Leenay, K.R. Maksimchuk, R.A. Slotkowski, R.N. Agrawal, A.A. Gomaa, A.E. Briner, et al., Identifying and visualizing functional PAM diversity across CRISPR–Cas systems, Mol. Cell 62 (2016) 137–147.
[46] J.N. Zadeh, C.D. Steenberg, J.S. Bois, B.R. Wolfe, M.B. Pierce, A.R. Khan, et al., NUPACK: analysis and design of nucleic acid systems, J. Comput. Chem. 32 (2011) 170–173.
[47] A.E. Briner, E.D. Henriksen, R. Barrangou, Prediction and validation of native and engineered Cas9 guide sequences, Cold Spring Harb Protoc 2016 (2016) (pdb prot086785).
[48] R. Sanozky-Dawes, K. Selle, S. O'Flaherty, T. Klaenhammer, R. Barrangou, Occurrence and activity of a type II CRISPR–Cas system in *Lactobacillus gasseri*, Microbiology 161 (2015) 1752–1761.
[49] K. Selle, T.R. Klaenhammer, R. Barrangou, CRISPR-based screening of genomic island excision events in bacteria, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 8076–8081.
[50] R.B. Vercoe, J.T. Chang, R.L. Dy, C. Taylor, T. Gristwood, J.S. Clulow, et al., Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands, PLoS Genet. 9 (2013), e1003454.
[51] J.H. Oh, J.P. van Pijkeren, CRISPR–Cas9-assisted recombineering in *Lactobacillus reuteri*, Nucleic Acids Res. 42 (2014), e131.
[52] K. Penewit, E.A. Holmes, K. McLean, M. Ren, A. Waalkes, S.J. Salipante, Efficient and scalable precision genome editing in *Staphylococcus aureus* through conditional recombineering and CRISPR/Cas9-mediated counterselection, MBio 9 (2018).
[53] A.A. Gomaa, H.E. Klumpe, M.L. Luo, K. Selle, R. Barrangou, C.L. Beisel, Programmable removal of bacterial strains by use of genome-targeting CRISPR–Cas systems, MBio 5 (2014), e00928-13.
[54] D. Bikard, C.W. Euler, W. Jiang, P.M. Nussenzweig, G.W. Goldberg, X. Duportet, et al., Exploiting CRISPR–Cas nucleases to produce sequence-specific antimicrobials, Nat. Biotechnol. 32 (2014) 1146–1150.
[55] R.J. Citorik, M. Mimee, T.K. Lu, Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases, Nat. Biotechnol. 32 (2014) 1141–1145.
[56] Y. Li, S. Pan, Y. Zhang, M. Ren, M. Feng, N. Peng, et al., Harnessing type I and type III CRISPR–Cas systems for genome editing, Nucleic Acids Res. 44 (2016), e34.
[57] F. Cheng, L. Gong, D. Zhao, H. Yang, J. Zhou, M. Li, et al., Harnessing the native type I-B CRISPR–Cas for genome

editing in a polyploid archaeon, J. Genet. Genomics 44 (2017) 541–548.
[58] A.E. Stachler, A. Marchfelder, Gene repression in haloarchaea using the CRISPR (clustered regularly interspaced short palindromic repeats)–Cas I-B system, J. Biol. Chem. 291 (2016) 15226–15242.
[59] Y. Chang, T. Su, Q. Qi, Q. Liang, Easy regulation of metabolic flux in *Escherichia coli* using an endogenous type I-E CRISPR–Cas system, Microb. Cell Factories 15 (2016) 195.
[60] M.L. Luo, A.S. Mullis, R.T. Leenay, C.L. Beisel, Repurposing endogenous type I CRISPR–Cas systems for programmable gene repression, Nucleic Acids Res. 43 (2015) 674–681.
[61] M.E. Pyne, M.R. Bruder, M. Moo-Young, D.A. Chung, C.P. Chou, Harnessing heterologous and endogenous CRISPR–Cas machineries for efficient markerless genome editing in *Clostridium*, Sci. Rep. 6 (2016) 25666.
[62] J. Zhang, W. Zong, W. Hong, Z.T. Zhang, Y. Wang, Exploiting endogenous CRISPR–Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production, Metab. Eng. 47 (2018) 49–59.
[63] B. Martel, S. Moineau, CRISPR–Cas: an efficient tool for genome engineering of virulent bacteriophages, Nucleic Acids Res. 42 (2014) 9504–9513.
[64] R. Kiro, D. Shitrit, U. Qimron, Efficient engineering of a bacteriophage genome using the type I-E CRISPR–Cas system, RNA Biol. 11 (2014) 42–44.
[65] W. Jiang, D. Bikard, D. Cox, F. Zhang, L.A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR–Cas systems, Nat. Biotechnol. 31 (2013) 233–239.
[66] Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, Multigene editing in the *Escherichia coli* genome via the CRISPR–Cas9 system, Appl. Environ. Microbiol. 81 (2015) 2506–2514.
[67] Y. Wang, Z.T. Zhang, S.O. Seo, P. Lynn, T. Lu, Y.S. Jin, et al., Bacterial genome editing with CRISPR–Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example, ACS Synth. Biol. 5 (2016) 721–732.
[68] F. Wasels, J. Jean-Marie, F. Collas, A.M. Lopez-Contreras, Ferreira N. Lopes, A two-plasmid inducible CRISPR/Cas9 genome editing tool for *Clostridium acetobutylicum*, J. Microbiol. Methods 140 (2017) 5–11.
[69] A.W. Westbrook, M. Moo-Young, C.P. Chou, Development of a CRISPR–Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*, Appl. Environ. Microbiol. 82 (2016) 4876–4895.
[70] J.P. Zhang, X.L. Li, A. Neises, W. Chen, L.P. Hu, G.Z. Ji, et al., Different effects of sgRNA length on CRISPR-mediated gene knockout efficiency, Sci. Rep. 6 (2016) 28566.
[71] M.L. Lemay, D.M. Tremblay, S. Moineau, Genome engineering of virulent lactococcal phages using CRISPR–Cas9, ACS Synth. Biol. 6 (2017) 1351–1358.
[72] J. Shen, J. Zhou, G.Q. Chen, Z.L. Xiu, Efficient genome engineering of a virulent *Klebsiella* bacteriophage using CRISPR–Cas9, J. Virol. 92 (2018) https://doi.org/10.1128/JVI.00534-18 pii: e00534-18.
[73] A.A. Dominguez, W.A. Lim, L.S. Qi, Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation, Nat. Rev. Mol. Cell Biol. 17 (2016) 5–15.
[74] T. Karvelis, G. Gasiunas, V. Siksnys, Harnessing the natural diversity and in vitro evolution of Cas9 to expand the genome editing toolbox, Curr. Opin. Microbiol. 37 (2017) 88–94.
[75] A. Mahas, C. Neal Stewart Jr., M.M. Mahfouz, Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation, Biotechnol. Adv. 36 (2018) 295–310.
[76] K. Li, D. Cai, Z. Wang, Z. He, S. Chen, Development of an efficient genome editing tool in *Bacillus licheniformis* using CRISPR–Cas9 Nickase, Appl. Environ. Microbiol. 84 (2018).
[77] Q. Li, J. Chen, N.P. Minton, Y. Zhang, Z. Wen, J. Liu, et al., CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*, Biotechnol. J. 11 (2016) 961–972.
[78] X. Song, H. Huang, Z. Xiong, L. Ai, S. Yang, CRISPR–Cas9 (D10A) Nickase-assisted genome editing in *Lactobacillus casei*, Appl. Environ. Microbiol. 83 (2017).
[79] K. Standage-Beier, Q. Zhang, X. Wang, Targeted large-scale deletion of bacterial genomes using CRISPR-Nickases, ACS Synth. Biol. 4 (2015) 1217–1225.
[80] T. Xu, Y. Li, Z. Shi, C.L. Hemme, Y. Li, Y. Zhu, et al., Efficient genome editing in *Clostridium cellulolyticum* via CRISPR–Cas9 Nickase, Appl. Environ. Microbiol. 81 (2015) 4423–4431.
[81] Y. Tong, P. Charusanti, L. Zhang, T. Weber, S.Y. Lee, CRISPR–Cas9 based engineering of actinomycetal genomes, ACS Synth. Biol. 4 (2015) 1020–1029.
[82] T. Su, F. Liu, P. Gu, H. Jin, Y. Chang, Q. Wang, et al., A CRISPR–Cas9 assisted non-homologous end-joining strategy for one-step engineering of bacterial genome, Sci. Rep. 6 (2016) 37895.
[83] X. Zheng, S.Y. Li, G.P. Zhao, J. Wang, An efficient system for deletion of large DNA fragments in *Escherichia coli* via introduction of both Cas9 and the non-homologous end joining system from *Mycobacterium smegmatis*, Biochem. Biophys. Res. Commun. 485 (2017) 768–774.
[84] L. Cui, D. Bikard, Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*, Nucleic Acids Res. 44 (2016) 4243–4251.
[85] A. Bernheim, A. Calvo-Villamanan, C. Basier, L. Cui, E.P.C. Rocha, M. Touchon, et al., Inhibition of NHEJ repair by type II-A CRISPR–Cas systems in bacteria, Nat. Commun. 8 (2017) 2094.