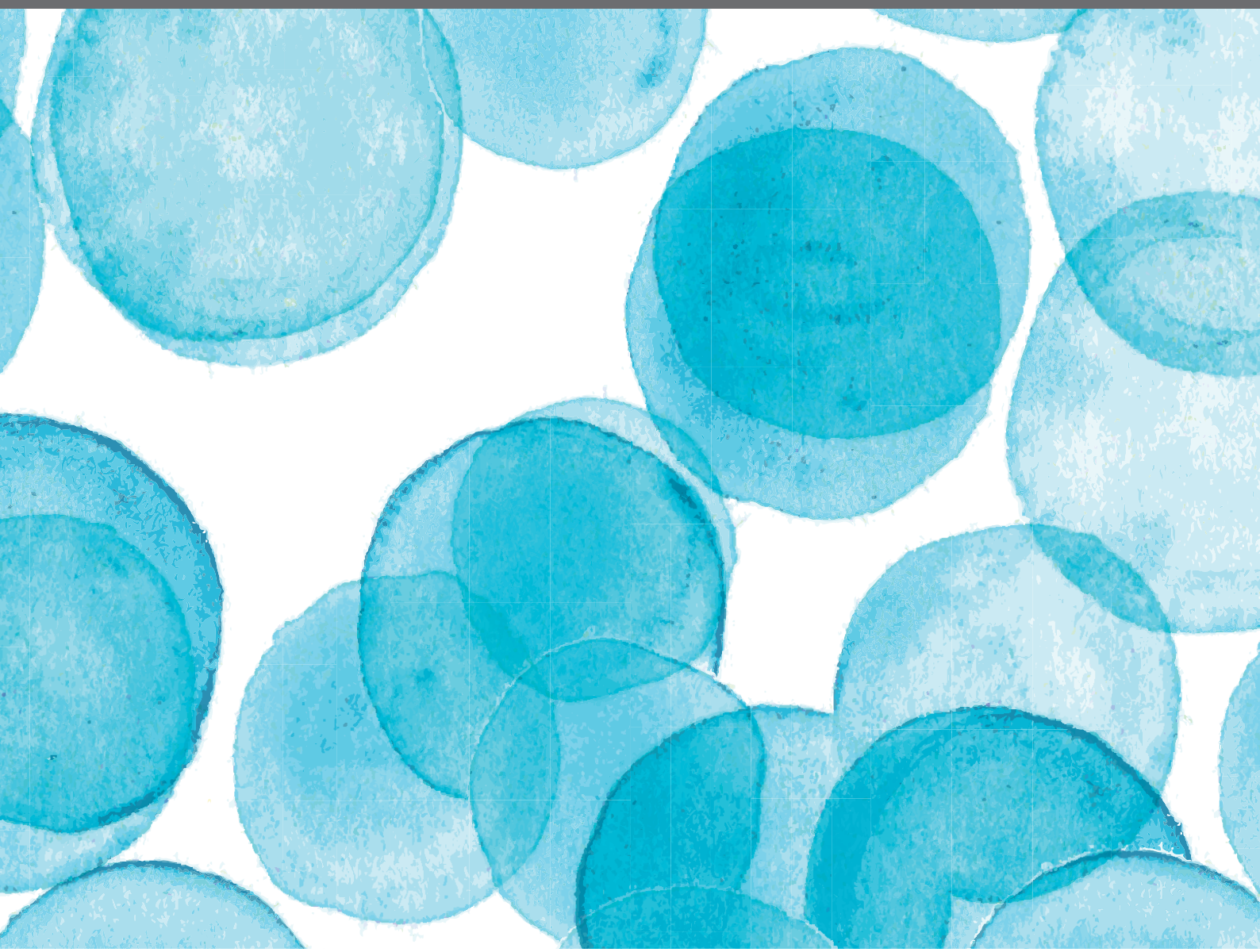
CRISPR-CAS SYSTEMS IN BACTERIA AND ARCHAEA

EDITED BY: Muhammad Kamruzzaman, Graciela Castro Escarpulli and Aixin Yan

PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-963-8
DOI 10.3389/978-2-88974-963-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

CRISPR-CAS SYSTEMS IN BACTERIA AND ARCHAEA

Topic Editors:
Muhammad Kamruzzaman, Westmead Institute for Medical Research, Australia
Graciela Castro Escarpulli, Instituto Politécnico Nacional de México (IPN), Mexico
Aixin Yan, The University of Hong Kong, Hong Kong, SAR China

Citation: Kamruzzaman, M., Escarpulli, G. C., Yan, A., eds. (2022). CRISPR-Cas Systems in Bacteria and Archaea. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88974-963-8

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 29 March 2022
doi: 10.3389/fmicb.2022.887778

Editorial: CRISPR-Cas Systems in Bacteria and Archaea

Muhammad Kamruzzaman[1]*, *Aixin Yan*[2] and *Graciela Castro-Escarpulli*[3]

[1] Centre for Infectious Diseases and Microbiology, Westmead Institute for Medical Research, The University of Sydney, Sydney, NSW, Australia, [2] School of Biological Sciences, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, [3] Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional de México (IPN), Mexico City, Mexico

Keywords: CRISPR-Cas, adaptive immunity, spacers, gene editing, bioethics

Editorial on the Research Topic

CRISPR-Cas Systems in Bacteria and Archaea

The CRISPR-Cas [clustered regularly interspaced short palindromic repeats and the CRISPR-associated genes (Cas)] is an adaptive immune system of prokaryotes against the invasion of foreign genetic elements and is widely distributed in the chromosomes of most archaea and many bacteria (Garneau et al., 2010; Marraffini, 2015; Hille et al., 2018). The system consists of a CRISPR array, comprising of short direct repeats, separated by short variable DNA sequences (called "spacers") acquired from foreign genetic elements and is flanked by various Cas genes. Cas genes are highly diverse and are involved in the different stages of CRISPR activity. Even though CRISPR-Cas is known as a defense system of prokaryotes, they are involved in different non-defense roles, including bacterial biofilm formation, regulation of quorum sensing, and pathogenicity. This special issue aims to collect articles that shed light on the recent advances in the CRISPR-Cas research to better understand the distribution, diversity, and biological functions of CRISPR-Cas systems. We have collected nine articles that highlight the recent studies on distribution, structure, biological functions and applications of CRISPR-Cas, as well as ethical considerations of CRISPR-Cas research.

Bioinformatic analysis of 716 genomes of *Staphylococcus aureus* (by Cruz-López et al.) identified that only 0.83% of *S. aureus* strains of the different geographical regions have type IIA CRISPR-Cas system, suggesting the occurrence of CRISPR-Cas in *S. aureus* may be spontaneous horizontal gene transfer event. 0.9% of the unique spacers matched with either plasmid or phage genomes, including bacteriophages used for the therapy against *S. aureus* infection, indicating the development of phage resistance *S. aureus* and therapeutic failure due to the CRISPR defense mechanism.

Direct uptake of foreign DNA from surrounding environments plays an important role in the genome diversity and evolution in bacteria and archaea. Liu et al. reviewed the functions and possible mechanism of the CRISPR systems and Argonauts in cellular defense against natural transformation. A limited number of studies demonstrated that type II CRISPR-Cas could prevent natural transformation in bacteria; however, the exact mechanism and whether other types of CRISPR systems also antagonize natural transformation is not known. Argonauts also can prevent the natural transformation of plasmid DNA. Unlike CRISPR-Cas systems, argonauts-mediated defenses do not integrate DNA fragments into host genomes and, thus, no memory of the invading DNA is generated.

To optimize sequence-specific immunity against invading genetic elements, CRISPR-Cas in prokaryotes continuously acquire spacers from the newly invading threats. Over time, many of the acquired spacers may become useless in their defense mechanism. Therefore, spacer uptake, their existence and loss must be regulated. A very interesting review by Garret compiled different observations and experimental designs to speculate

OPEN ACCESS

Edited and reviewed by:
Marc Strous,
University of Calgary, Canada

***Correspondence:**
Muhammad Kamruzzaman
muhammad.kamruzzaman@sydney.edu.au

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *02 March 2022*
Accepted: *11 March 2022*
Published: *29 March 2022*

Citation:
Kamruzzaman M, Yan A and Castro-Escarpulli G (2022) Editorial: CRISPR-Cas Systems in Bacteria and Archaea. Front. Microbiol. 13:887778. doi: 10.3389/fmicb.2022.887778

a model for the spacer dynamics in the CRISPR array and demonstrated that new spacers are added at the leader end of the array, which varies among species, systems, and conditions. Rearrangement of the array is ongoing at some level, though the particular frequency is also variable among species and CRISPR-Cas classes. The terminal spacer-repeat unit rarely participates in rearrangements, so the array is maintained, and the last spacer-repeat unit is stable.

Type IV CRISPR-Cas system, primarily found on plasmids (Kamruzzaman and Iredell, 2019), is least understood among the six CRISPR types. The lack of Cas nucleases, integrases, and other genetic features commonly found in most CRISPR systems has made it difficult to predict the mechanisms of action and biological functions of type IV CRISPR-Cas. The perspective by Taylor et al. compiled and analyzed recent advances in bioinformatics, biochemical, and structural studies of type IV systems that provided valuable insights to understand the structure and function of type IV systems. Instead of Cas gene *csf1*, Cas-7 like gene *csf2* was proposed to be employed to distinguish type IV from other types in the Class1 CRISPR-Cas group. Type IV-A systems protect bacteria from plasmids and phages, which needs DinG helicase along with other Cas proteins with an unknown mechanism of action. Recently identified type IV-C systems lack a Csf1 subunit and instead encode a Cas10-like subunit with an HD nuclease Domain, while type IV-B systems lack a CRISPR locus and a crRNA processing enzyme and are associated with an ancillary gene identified as cysH-like. The mechanism of action and biological functions of type IV-B and -C are yet unknown.

Type III CRISPR-Cas systems can target both RNA and single-stranded DNA and provide immunity against invaders, which is dependent on the target RNA transcription. The target RNA binding also activates the cyclic oligoadenylate (cOA) synthesis activity of Cas10 subunit. The recent advances on cOA synthesis, cOA-activated effector protein, cOA signaling-mediated immunoprotection, cOA signaling inhibition, and possible crosstalk between cOA signaling and other cyclic oligonucleotide-mediated immunity have been discussed in the review by Huang and Zhu.

CRISPR-Cas is not only involved in bacterial defense mechanisms but also involved in the regulation of bacterial physiology. One example is the very conserved CRISPR-Cas system found in the *Salmonella* Typhi, which regulates the synthesis of major outer membrane proteins (OMP) OmpC, OmpF, OmpA and quiescent OMP, OmpS2 by regulating the expression of the master porin regulator OmpR (Medina-Aparicio et al.). This CRISPR-Cas system is also involved in the resistance to bile salts and in the formation of biofilms by *Salmonella* Typhi.

The application of CRISPR-Cas in genome editing for both prokaryotes and eukaryotes revolutionized genetic engineering technology. CRISPR-Cas gene-editing tool has enormous potential as antimicrobial agents, and Wang et al. successfully eliminated two virulence plasmids from *Bacillus anthracis* and *B. cereus* and specifically killed *B. anthracis* using the CRISPR/Cas9 system. The nuclease activity of CRISPR-Cas protein allows researchers to edit a genome with unprecedented ease, accuracy, and high throughput, while CRISPR interference (CRISPRi) technology has been developed for silencing specific genes by exploiting the catalytically inactive Cas9 (dCas9) and single-guide RNA (sgRNA). RNA interference (RNAi) technology is mainly used in eukaryotes to investigate the function of essential genes. The development of the CRISPRi system will provide a high-throughput, practical, and efficient tool for the discovery of functionally important genes in bacteria. The mini-review by Zhang et al. discussed the CRISPRi system, the underlying mechanism and properties and highlighted its application as a high-throughput screening tool in gene function analysis.

Finally, the precise gene editing capacity of CRISPR-Cas opens new possibilities to treat genetic diseases which are untreatable so far. But cautions need to be taken, and the processes need to be regulated to ensure patients' safety and implementation of bioethics (review by Gonzalez-Avila et al.). Scientists first need to understand the risk associated with any particular genetic modification, and CRISPR-based therapeutics must not be misprescribed or used for personal prejudices but always be approved by institutional and specialized bioethics committees. Scientists and world leaders should set boundaries about the use of CRISPR technologies.

AUTHOR CONTRIBUTIONS

MK prepared the call for this Research Topic. MK, AY, and GC-E edited the manuscript that were submitted. MK wrote this editorial. AY and GC-E reviewed it. All authors approved the submitted version.

FUNDING

MK was supported by the Australian National Health and Medical Research Council (NHMRC) Grant (APP1197534) during the writing of this editorial.

ACKNOWLEDGMENTS

MK, AY, and GC-E would like to thank all authors, reviewers, and editors for their contribution to this Research Topic.

REFERENCES

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

Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Kamruzzaman, M., and Iredell, J. R. (2019). CRISPR-Cas system in antibiotic resistance plasmids in *Klebsiella pneumoniae*. *Front Microbiol* 10, 2934. doi: 10.3389/fmicb.2019.02934

Marraffini, L. A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61. doi: 10.1038/nature15386

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

REVIEW
published: 22 October 2020
doi: 10.3389/fmicb.2020.593301

The Clustered Regularly Interspaced Short Palindromic Repeat System and Argonaute: An Emerging Bacterial Immunity System for Defense Against Natural Transformation?

Mafeng Liu[1,2,3]†, Mi Huang[1,2,3]†, Mingshu Wang[1,2,3], Dekang Zhu[2,3], Renyong Jia[1,2,3], Shun Chen[1,2,3], Ling Zhang[2,3], Leichang Pan[2,3] and Anchun Cheng[1,2,3]**

[1]Institute of Preventive Veterinary Medicine, Sichuan Agricultural University, Chengdu, China, [2]Research Centre of Avian Disease, College of Veterinary Medicine of Sichuan Agricultural University, Chengdu, China, [3]Key Laboratory of Animal Disease and Human Health of Sichuan Province, Sichuan Agricultural University, Chengdu, China

Clustered regularly interspaced short palindromic repeat (CRISPR) systems and prokaryotic Argonaute proteins (Agos) have been shown to defend bacterial and archaeal cells against invading nucleic acids. Indeed, they are important elements for inhibiting horizontal gene transfer between bacterial and archaeal cells. The CRISPR system employs an RNA-guide complex to target invading DNA or RNA, while Agos target DNA using single stranded DNA or RNA as guides. Thus, the CRISPR and Agos systems defend against exogenous nucleic acids by different mechanisms. It is not fully understood how antagonization of these systems occurs during natural transformation, wherein exogenous DNA enters a host cell as single stranded DNA and is then integrated into the host genome. In this review, we discuss the functions and mechanisms of the CRISPR system and Agos in cellular defense against natural transformation.

Keywords: Clustered regularly interspaced short palindromic repeat-Cas, Argonaute proteins, bacterial immunity system, natural transformation, ssDNA

OPEN ACCESS

Edited by:
Graciela Castro Escarpulli, National Polytechnic Institute of Mexico (IPN), Mexico

Reviewed by:
Muhammad Zubair Shabbir, University of Veterinary and Animal Sciences, Pakistan
Yinhua Lu, Shanghai Normal University, China

***Correspondence:**
Anchun Cheng
chenganchun@vip.163.com
Mafeng Liu
liumafengra@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *11 August 2020*
Accepted: *23 September 2020*
Published: *22 October 2020*

Citation:
Liu M, Huang M, Wang M, Zhu D, Jia R, Chen S, Zhang L, Pan L and Cheng A (2020) The Clustered Regularly Interspaced Short Palindromic Repeat System and Argonaute: An Emerging Bacterial Immunity System for Defense Against Natural Transformation? Front. Microbiol. 11:593301. doi: 10.3389/fmicb.2020.593301

INTRODUCTION

Horizontal exchange of DNA between bacteria is an important mechanism to generate genome diversity and drive evolution (Gogarten and Townsend, 2005). For example, the emergence of super resistant, virulent bacterial strains has largely been inferred to be caused by the transfer of antimicrobial resistance and virulence genes among different species (Maeusli et al., 2020). The acquisition of genetic material can occur through natural transformation, direct DNA uptake from the environment, conjugation, plasmid transfer from other cells, and transduction, the latter of which incorporates heterologous DNA from bacteriophage infection (Arber, 2014; Darmon and Leach, 2014). Among these, only natural transformation is exclusively facilitated by genes from the bacterial chromosome (Ambur et al., 2016).

Traditionally, the active acquisition of genetic material has been thought to benefit the recipient bacterium. For example, *Streptococcus pneumoniae* required exogenous DNA to become antibiotic resistant and inhibit vaccination treatments (Croucher et al., 2011). However, it was recently suggested that this conclusion was biased because only bacterial genomes that survived selection were ultimately observed as recipients (Ambur et al., 2016). Further, bacteria have been observed to uptake fragmented and damaged DNA by natural transformation when the DNA contains abasic sites or miscoding lesions (Thomas and Nielsen, 2005; Overballe-Petersen et al., 2013). Thus, bacteria indiscriminately take up both "beneficial" and "harmful" DNA. Moreover, these newly integrated gene elements can also alter the structure of the recipient genome and introduce additional physiological burdens (Johnston et al., 2014; Blokesch, 2017).

To control the entry of new genetic material, bacteria have developed immune defenses to limit inter-specific or inter-strain horizontal transfer of chromosomal DNA. Bacterial immune systems, including restriction-modification (R-M) systems, CRISPR systems, and Argonaute proteins (Agos), have recently been observed to play key roles in defending bacterial cells from intrusion of foreign DNA including *via* bacteriophages and plasmid DNA (Bikard et al., 2012; Zhang and Blaser, 2012; Swarts et al., 2014a). However, natural transformation internalizes exogenous ssDNA and integrates it into the host chromosome by homologous recombination. This implies that these immune system mechanisms antagonize natural transformation by unique mechanisms or functions at different stages of natural transformation. For example, R-M systems have been proposed to target natural transformation after the replication of integrated ssDNA into a host chromosome (Johnston et al., 2013).

CRISPR systems were recently shown to inhibit natural transformation in *S. pneumoniae* and *Neisseria meningitidis* (Bikard et al., 2012; Zhang et al., 2013), while the Agos of *Thermus thermophilus* have been shown to prevent the uptake and propagation of naturally transformed plasmid DNA (Swarts et al., 2014a). However, the mechanisms by which CRISPRs and Agos prevent natural transformation has not been fully described in these studies.

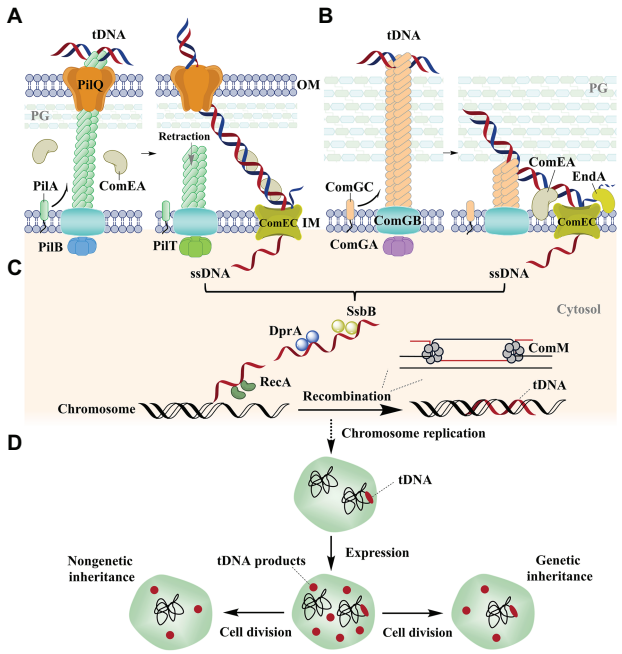


FIGURE 1 | Natural transformation in the Gram-negative bacterium *Vibrio cholerae* (Seitz and Blokesch, 2013; Seitz et al., 2014; Ellison et al., 2018) and the Gram-positive bacterium *Streptococcus pneumoniae* (Laurenceau et al., 2013, 2015; Balaban et al., 2014; Muschiol et al., 2015). **(A)** Natural transformation of *V. cholerae*. Extracellular DNA is bound to the pilus and transported across the OM with the participation of the periplasmic protein ComEA. The pilus is composed of PilA (the main fiber subunit), the PilB and PilT ATPases, PilQ (a secretin pore), PliC (the motor), and other proteins (Matthey and Blokesch, 2016; McCallum et al., 2019). **(B)** Natural transformation of *S. pneumoniae*. DNA is bound to a pseudopilus and is transported across the cell wall with the participation of the membrane protein ComEA. The pseudopilus is mostly composed of ComGC (the major pilin), ComGB (a membrane protein), and ComGA (an ATPase; Muschiol et al., 2015). One strand is degraded by the EndA nuclease, while the other strand is transported into the cytoplasm by ComEC, which also occurs in *V. cholerae*. **(C)** ssDNA integration into the genome. DprA binds internalized ssDNA and recruits the recombinase RecA that integrates ssDNA into the host genome with the help of ComM. **(D)** The replication and expression of the integrated DNA in the cell. After replication of the genome, the host cell contains two genome types. The integrated DNA is immediately expressed in the host cell and the expression product is propagated by both non-genetic and genetic inheritance mechanisms after cell division. tDNA, transforming DNA; M, membrane; OM, outer membrane; IM, inner membrane; PG, peptidoglycan.

AN OVERVIEW OF THE NATURAL TRANSFORMATION PROCESS

Bacterial natural transformation is a complex process involving uptake of extracellular DNA to the cytoplasm and integration into the chromosome. Based on *Helicobacter pylori* (Stingl et al., 2010), *Neisseria* (Maier et al., 2002; Gangel et al., 2014), and *Vibrio cholerae* (Seitz and Blokesch, 2013; Seitz et al., 2014; Ellison et al., 2018) models, Gram-negative bacterial DNA uptake requires its transport across the outer membrane and the translocation of DNA across the inner membrane (**Figure 1A**).

In natural transformation by *V. cholerae*, the type IV family of pili (T4P) on the surface of most Gram-negative bacteria retracts DNA into the periplasm through direct binding of the extracellular double-stranded DNA (dsDNA) on their ends (Chen and Dubnau, 2004; Claverys et al., 2009; Burton and Dubnau, 2010; Ellison et al., 2018; **Figure 1A**). Nevertheless, the method by which pili specifically bind DNA remains unclear. It was recently shown that ComE(A), a periplasmic DNA-binding protein, is essential for the uptake and transport of DNA from the outer membrane to the periplasm (Gangel et al., 2014; Hepp and Maier, 2016; **Figure 1A**). Further, T4P and ComE(A) mediate uptake *via* a Brownian ratchet mechanism (Hepp and Maier, 2016; Dubnau and Blokesch, 2019). An exception to this mechanism has been observed in *H. pylori*, which employs a ComB type-IV secretion system (T4SS), rather than the T4P system, for initial DNA uptake during transformation (Karnholz et al., 2006). ComH is a periplasmic DNA-binding

protein that is involved in the transport of DNA into the periplasm, although the interaction between ComB and ComH remains unknown (Damke et al., 2019).

Gram-positive bacteria like *S. pneumonia* use pseudopili to transport extracellular DNA through a thick layer of peptidoglycan (Laurenceau et al., 2013, 2015; Balaban et al., 2014; Muschiol et al., 2015; **Figure 1B**). In contrast, the Gram-positive bacterium *Bacillus subtilis* initiates DNA binding independent of a pseudopilus (Hahn et al., 2005; Kidane and Graumann, 2005; Mirouze et al., 2018). Rather, it was recently shown that wall teichoic acids (WTAs) are responsible for the initial step in transformation (Mirouze et al., 2018). After the DNA is in the periplasm of Gram-negative cells or in the compartments between the cell wall and the membrane of Gram-positive bacterial cells, one strand is degraded to nucleotides at the membrane surface and the other is internalized into the cytoplasm in single-stranded form through the ComEC transmembrane channel (Draskovic and Dubnau, 2005; Mell and Redfield, 2014; **Figures 1A,B**). However, the protein that degrades the non-transforming DNA and the mechanism that is used for degradation remains unknown. The internalized single-stranded DNA (ssDNA) is then bound by DNA processing protein A (DprA) that recruits the recombinase RecA to the ssDNA (Mortier-Barriere et al., 2007) and the translocated strand can then be used to replace a chromosomal strand *via* recombination (**Figure 1C**). During recombination, RecA is responsible for identifying homologous DNA regions and initiates strand invasion to form a displacement loop (D-loop) in the chromosome (**Figure 1C**). ComM subsequently promotes expansion of the D-loop using a bidirectional helicase and branch migration activities that enhance the integration efficiency of ssDNA into the genome (Nero et al., 2018; **Figure 1C**). The biological functions and mechanisms of action for ComE(A), ComEC, DprA, RecA, and ComM are all evolutionarily conserved in most competent bacteria (Provvedi and Dubnau, 1999; Berge et al., 2002; Draskovic and Dubnau, 2005; Berge et al., 2013; Johnston et al., 2013; Johnston et al., 2014; Seitz et al., 2014; Hepp and Maier, 2016; Salzer et al., 2016; Nero et al., 2018; Pimentel and Zhang, 2018; Huang et al., 2019).

Following these activities, the integrated DNA is immediately expressed after chromosomal replication, which has been demonstrated to occur before cell division in *V. cholera* (Dalia and Dalia, 2019), *H. pylori* (Corbinais et al., 2016), and *B. subtilis* (Boonstra et al., 2018; **Figure 1D**). Thus, the expressed products of integrated DNA can be used by the untransformed relatives of transformed cells, which are termed the non-genetic inheritance of traits (Dalia and Dalia, 2019; **Figure 1D**).

THE CRISPR SYSTEM AND NATURAL TRANSFORMATION

An Overview of CRISPR Systems

Jansen et al. (2002) first discovered clustered regularly interspaced short palindromic repeats (CRISPR) loci and the CRISPR-associated Cas genes. The genomes of ∼50% of bacteria and ∼90% of archaea have a CRISPR-Cas system that plays a role in defense against the inclusion of foreign (e.g., phage or plasmid) DNA (Makarova and Koonin, 2015; Makarova et al., 2015; Samson et al., 2015). CRISPR loci are composed of ∼24–48 bp short repetitive sequence arrays that are separated by equally short "spacer" sequences that are derived from mobile genetic elements like bacteriophages and plasmid sequences (Mojica et al., 2005). CRISPR-Cas systems have been divided into two distinct classes: class 1 and class 2. Class 1 systems have multi-subunit effector complexes, whereas class 2 systems have individual single-protein effector modules (Koonin et al., 2017; Koonin and Makarova, 2019). Furthermore, class 1 systems contain three different sub-types (types I, III, and IV), while class 2 systems contain types II, V, and VI. Each sub-type is characterized by distinct effector module architectures that contain unique signature proteins, like Cas3 for the type I systems, Cas9 for type II systems, and Cas10 for type III systems (Makarova and Koonin, 2015; Koonin et al., 2017; Koonin and Makarova, 2019).

CRISPR-Cas systems operate through three steps. In the first (the adaptation phase), the acquisition of spacers is often derived from phage or foreign plasmid DNA (Mojica et al., 2005). Importantly, the 2–5 nt protospacer adjacent motif (PAM) in the invading DNA is required for spacer acquisition (Marraffini and Sontheimer, 2010a). During adaptation, the Cas1-Cas2 complex is also required to process foreign DNA and subsequent integration into a CRISPR array (Hille et al., 2018). In the second step, short CRISPR RNAs (crRNAs) undergo biogenesis and maturation. In most bacteria, the repeat/spacer arrays are transcribed as long CRISPR RNA precursors (pre-crRNA) that are then cleaved within the repeat sequences and transformed into small crRNAs by Cas endoribonucleases (Hatoum-Aslan et al., 2011), which then base pair with foreign DNA *via* spacer-encoded sequences (Brouns et al., 2008). In class 1 systems, Cas-6-family enzymes are involved in processing RNA into mature crRNAs, while Cas9 is involved in the maturation of crRNAs in class 2, type II systems (Hille et al., 2018). The third step is the interference phase, wherein crRNAs are used as antisense guides that combine with sets of Cas proteins to form the core CRISPR-Cas ribonucleoprotein complexes. After complementary "protospacer" sequences from foreign invading DNA or RNA are recognized by these complexes, they are cleaved *via* sequence-specific mechanisms (Garneau et al., 2010). The PAM of invading DNA is also required for efficient CRISPR interference (Fischer et al., 2012). Further, crRNA-Cas complexes are base-paired with PAM sequences to avoid autoimmunity (Marraffini and Sontheimer, 2010b). Thus, type I and type II CRISPR-Cas systems cleave DNA, while type III systems can cleave DNA or RNA molecules (Barrangou and Marraffini, 2014). Specifically, *cas9* of type II CRISPR-Cas systems encodes a multidomain protein that contains all the functions of effector complexes and targets DNA cleavage sites (Jinek et al., 2012). Based on the above observations, the CRISPR system has been described as an RNA-guided "adaptive immune system" of bacteria. In addition, CRISPR-Cas systems also exhibit off-target activities on sequences that are similar to target sequences (Nivala et al., 2018), which may represent a potential "escape" mechanism for foreign DNA.

The CRISPR System Antagonizes Natural Transformation

Numerous studies have demonstrated that CRISPR systems function in defense against bacteriophage infection and plasmid transformation (Garneau et al., 2010; Amitai and Sorek, 2016). Accordingly, most archaeal spacers correspond to plasmids or bacteriophages (Brodt et al., 2011). Further, some of these sequences can be mapped to the chromosomal genes of other archaea, suggesting that CRISPR/Cas systems may also be involved in reducing the intrusion of foreign chromosomal DNA *via* natural transformation (Brodt et al., 2011). Bikard et al. (2012) introduced the CRISPR01 locus of *Streptococcus pyogenes* SF370 into the non-encapsulated strain, *S. pneumoniae* R6. The chromosomal DNA of the *S. pneumoniae* strain carrying the engineered spacer 1 (*spc1*) target was used as the donor for transformation with *S. pneumoniae* R6 and the *S. pneumoniae* crR6 strains. The *S. pneumoniae* crR6 strain prevented DNA transformation, while the control *S. pneumoniae* R6 strain could not prevent transformation (Bikard et al., 2012). Furthermore, a spacer for the *cap* gene was introduced into the CRISPR locus of *S. pneumoniae* crR6, and CRISPR interference was able to prevent capsule-switching of pneumococci both *in vitro* and *in vivo* (Bikard et al., 2012). Thus, CRISPR/Cas systems can prevent natural transformation, at least in *S. pneumoniae*. Nevertheless, it is unclear if native CRISPR/Cas systems can limit natural transformation, since all known published pneumococcal genomes do not encode CRISPR loci (Makarova et al., 2011).

Zhang et al. (2013) subsequently cloned protospacers with flanking *Neisseria* sequences that conform to the PAM consensus sequence into the pGCC2 plasmid. It was then shown that the plasmid pGCC2 can be integrated into the genome of *N. meningitidis* through natural transformation. However, a plasmid containing protospacers failed to integrate into *N. meningitidis* through natural transformation, in contrast to the empty plasmid (Zhang et al., 2013). Mutation of two consecutive nucleotides within the PAM or seed sequence of the protospacer in the plasmid abolished CRISPR interference, thereby reestablishing the natural transformation potential. In contrast, nucleotide substitutions in a non-PAM flanking region or substitutions in non-seed protospacer positions did not affect CRISPR interference (Zhang et al., 2013). Similarly, natural transformation was inhibited when using genomic DNA containing protospacers. The CRISPR-Cas systems in both of the above studies were type II systems that use Cas9 proteins as effector proteins to cleave invading nucleic acids (Makarova et al., 2015).

Evolutionary Analysis of Natural Transformation and CRISPR Systems

The evolutionary association between natural transformation and CRISPR systems has been investigated through comparative genomics. The loss of competence in *Aggregatibacter actinomycetemcomitans* was strongly and positively correlated with the loss of a CRISPR system (Jorth and Whiteley, 2012). Moreover, the genomes of competent bacteria were larger and contained multiple rearrangements in contrast to the non-competent strain genomes. Rather, non-competent bacterial genomes were extremely stable, but susceptible to infective DNA element integration (Jorth and Whiteley, 2012), suggesting that CRISPRs play a role in defense against exogenous DNA invasion.

Using the length of a CRISPR array as a proxy for CRISPR activity, Gophna et al. (2015) analyzed the connection between CRISPR activity and gene acquisition *via* horizontal gene transfer (HGT). CRISPR-negative bacterial genomes encoded fewer prophage-encoded proteins on average compared to CRISPR-positive genomes, suggesting that CRISPR systems do not inhibit HGT on evolutionary timescales (Gophna et al., 2015). It was instead hypothesized that the resistance of CRISPR-Cas systems to mobile elements occurs at the population scale rather than over evolutionary timescales (Gophna et al., 2015).

The Csx27 protein of subtype VI-B1 CRISPR-Cas systems was recently shown to be encoded in the same predicted operons as the components of natural transformation systems (Makarova et al., 2019). The Csx27 protein has four predicted transmembrane regions, and it was thus predicted that Csx27 proteins form membrane channels for the transport or degradation of ssDNA (Makarova et al., 2019). Taken together, the interactive relationships between natural transformation and CRISPR systems cannot be adequately established through evolutionary analysis. Nevertheless, the connection between CRISPRs and natural transformation requires further experimental investigation.

The Unstated Conundrum and a Proposed Model for CRISPR System Limiting of Natural Transformation

Although the above two studies suggested that CRISPR systems can limit natural transformation, several associated conundrums have not been resolved. First, it is unclear whether and how new spacers are acquired from internalized ssDNA. Second, it is unknown if CRISPR systems target internalized ssDNA or dsDNA after exogenous genes are integrated and replicated into the host genome. To our knowledge, no reports indicate that CRISPR systems can acquire spacers from ssDNA. The RNA guided Cas9 proteins from *N. meningitidis* and *S. pyogenes* were recently shown to cleave ssDNA in a PAM- and tracrRNA-independent manner (Ma et al., 2015; Zhang et al., 2015), suggesting that the CRISPR systems of *N. meningitidis* and *S. pyogenes* target ssDNA. In addition, Cas12a of Lachnospiraceae and the Cas1 protein of *Escherichia coli* (Babu et al., 2011) have been observed to target ssDNA and degrade it (Chen et al., 2018). A model was proposed for the antagonization of heterologous transformation by CRISPR systems in combination with the proposed post-replication targeting model for CRISPR systems (Johnston et al., 2013) using the *Neisseria* CRISPR system as an example (**Figure 2**). In the model, (I) spacers are pre-acquired from internalized ssDNA through an unidentified pattern and (II) heterologous dsDNA enters the cytoplasm as ssDNA through natural transformation. About 50% of ssDNA is targeted and cleaved by the crRNA and Cas9 endonuclease complexes in this stage, since each strand of the dsDNA randomly enters the cell. Thus, a portion of ssDNA escapes this immune response, and (III) after integration and replication, the double-stranded DNA associated with a PAM is produced in the transformant chromosome.

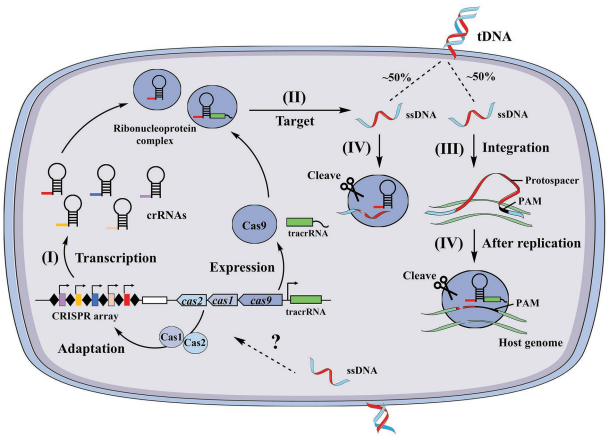


FIGURE 2 | A proposed model of type II clustered regularly interspaced short palindromic repeat (CRISPR) system inhibition of natural transformation. CRISPRs consist of short repeats separated by non-repetitive spacers that are obtained from previous invading sequences. (I) CRISPR loci are transcribed as long CRISPR RNA (crRNA) precursors and then cleaved into small mature crRNAs, (II) during natural transformation, heterologous DNA enters the cell as ssDNA that is then targeted by the CRISPR system, (III) after replication in the transformant chromosome, a dsDNA that contains a "protospacer" and PAM sequence is produced, and (IV) the complex of crRNAs and Cas proteins recognize and cleave the complementary "protospacer" in the transformed chromosomes.

At this stage, the CRISPR system in the host genome can target and cleave the genomic PAM site, but the CRISPR system cannot target itself, as these regions lack a PAM site (Johnston et al., 2013; **Figure 2**). In support of this hypothetical model, the transfer of active CRISPR/Cas systems into a recipient cell containing a target sequence has been shown to result in cell death (Bikard et al., 2012). This model is based on the requirement of a spacer to pre-exist in the host genome and coincide with the protospacers of invading DNA. Nevertheless, the fundamental question as to whether native CRISPR/Cas systems acquire new spacers from internalized ssDNA remains unaddressed. Another important question that remains to be answered is whether other types of CRISPR/Cas systems are involved in antagonizing natural transformation.

ARGONAUTES INHIBIT NATURAL TRANSFORMATION

Argonaute proteins (Agos) were initially discovered in eukaryotes and were later observed as conserved across all domains of life (Bohmert et al., 1998; Hock and Meister, 2008). Agos bind small noncoding RNAs in eukaryotes and target complementary RNA to regulate gene expression and repress invasive genomic elements (Peters and Meister, 2007). The homologues of Argonautes are present in some bacterial (~9%) and archaeal (~32%) genomes (Makarova et al., 2009; Swarts et al., 2014b). However, bacterial and archaeal Agos do not encode Dicer homologs and the TAR RNA-binding protein (TRBP) that is important for the silencing pathway.

Argonaute Interference With the Replication of Exogenous Plasmids Internalized by Natural Transformation

The domain organization of some prokaryotic Argonautes was observed to be similar to eukaryotic orthologs (Song et al., 2004; Yuan et al., 2005; Wang et al., 2008). However, archaeal and bacterial Argonautes have a higher affinity for ssDNA and dsDNA compared to eukaryotic homologs (Ma et al., 2005; Yuan et al., 2005). Nevertheless, experimental evidence to understand the functions of prokaryotic Argonautes in host defenses have not appeared until recently. *ago* mutation in *T. thermophilus* leads to increased natural transformation efficiency of plasmids by 10-fold compared to wild-type cells (Swarts et al., 2014a). Moreover, plasmid yields from wild-type cells are lower than those of *ago* knockout strain (Swarts et al., 2014a). Analysis of co-purified nucleic acids revealed that the Ago from *T. thermophilus* (TtAgo) binds 13–25 bp ssDNAs that are mostly derived from plasmids and have a strong bias for the 5′-end deoxycytidine (Swarts et al., 2014a). Plasmid cleavage assays also showed that guide DNA-loaded TtAgo was able to cleave both single- and double-stranded targets (Swarts et al., 2014a). Thus, the authors speculated that TtAgo uses siDNA guides to specifically cleave ssDNA targets that are produced during natural transformation (Swarts et al., 2014a; **Figure 3**). Recent studies have suggested that most characterized Argonautes from bacteria and archaea function to target complementary dsDNA or ssDNA against invasive genetic elements, in contrast to their functions in eukaryotes (Makarova et al., 2009; Olovnikov et al., 2013; Swarts et al., 2014a). Agos bind small RNAs (15–19 nt) in *Rhodobacter sphaeroides* that are derived from mRNAs or are the products of their degradation (Olovnikov et al., 2013). In addition, Olovnikov et al. (2013) observed strong degradation of a plasmid upon the expression of RsAgo in *E. coli* cells, although Argonaute-dependent cleavage activity was not detected. Thus, it was proposed that RsAgo use RNA guides to recruit an associated nuclease for subsequent target cleavage (Olovnikov et al., 2013).

The DNA Cleavage Mechanism of Bacterial and Archaeal Argonautes

A remaining question was how guiding DNAs were generated and loaded onto Agos in bacterial cells. Guide-free TtAgo, SeAgo, and MjAgo were shown to degrade double-stranded plasmid and genomic DNA to 8–26 nt oligonucleotides, an activity that was termed "DNA chopping," with the subsequent small dsDNA fragments loaded onto Agos (Swarts et al., 2017; Zander et al., 2017; Olina et al., 2020). Moreover, the cleavage efficiency of pre-loaded MjAgo for a plasmid was higher when compared to apo-MjAgo (Zander et al., 2017). Similarly, the use of cleavage products as functional guides was also demonstrated for TtAgo (Swarts et al., 2017). In addition to TtAgo, several bacterial and archaeal Argonautes have been shown *in vitro* to cleave target DNA using ssDNA as a guide, including the Argonaute from *Pyrococcus furiosus* (PfAgo; Swarts et al., 2015), the Argonaute of *Clostridium butyricum* (CbAgo; Hegge et al., 2019; Kuzmenko et al., 2019), the Argonaute of

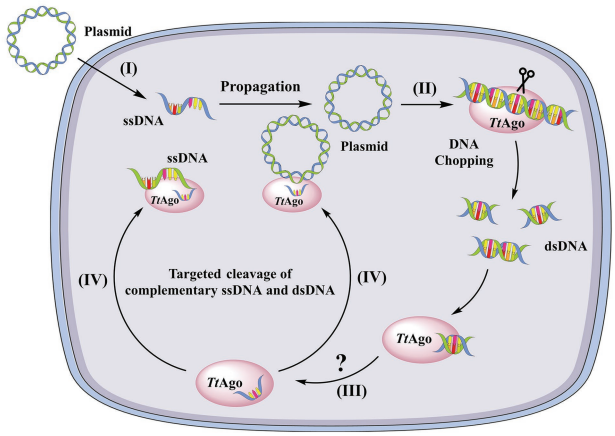


FIGURE 3 | A model of TtAgo-mediated DNA-guided plasmid inhibition. (I) a plasmid enters the cell as ssDNA through natural transformation. At this stage, the ssDNA is resistant to apo-TtAgo, (II) after replication, the dsDNA plasmid is chopped by apo-TtAgo, (III) the fragmented dsDNA is then bound by TtAgo and one strand of the DNA is released from TtAgo through an unknown mechanism, and (IV) the siDNA-loaded TtAgo targets complementary DNA (ssDNA and dsDNA) and cleaves it.

Limnothrix rosea (LrAgo; Kuzmenko et al., 2019), the Argonaute of *Methanocaldococcus jannaschii* (MjAgo; Zander et al., 2017) and the Argonaute of *Synechococcus elongatus* (SeAgo; Olina et al., 2020). Additionally, the Argonautes in some bacteria, such as *Marinitoga piezophile*, *Thermotoga profunda*, and *Rhodobacter sphaeroides*, use RNA as a guide to target DNA (Kaya et al., 2016; Miyoshi et al., 2016).

A Proposed Model and Remaining Conundrums of the Inhibition of Natural Transformation by Argonautes

In summary, a scenario can be envisioned wherein pAgo, using TtAgo as an example, limits natural transformation when a plasmid is the substrate DNA, as shown in **Figure 3**. In this generalized mechanism, (I) a plasmid enters a cell by natural transformation as ssDNA that is resistant to the "chopping activity" of TtAgo since the guide-free TtAgo cannot degrade ssDNA (Swarts et al., 2017), (II) after propagation, the plasmid DNA is chopped by apo-TtAgo and the small DNA fragments are loaded into TtAgo (**Figure 3**), and (III) the target strand is then dissociated from TtAgo through unknown mechanisms and TtAgo-siDNA attacks complementary ssDNA and dsDNA (**Figure 3**). However, unanswered questions remain regarding these mechanisms. First, if TtAgo has both "DNA chopping" and ssDNA guided cleavage activities, then all transformed plasmids should be cleaved. However, un-cleaved plasmids remain within *T. thermophilus* (Swarts et al., 2014a). Second, it is unknown how endogenous plasmids coexist with TtAgo and also how small DNA-loaded Ago complexes distinguish target foreign DNA from normal genomic substrates. Even over-expressed Ago in *S. elongatus* did not affect cell growth (Olina et al., 2020), suggesting that Ago does not disrupt the genome of *S. elongatus*. Further, it has been proposed that *M. jannaschii* histones, but not methylation, protects genomic DNA from Ago chopping activities (Zander et al., 2017). However, the *T. thermophilus* genome does not encode histones (Willkomm et al., 2018). An additional question is how one of the DNA strands released from Ago binds to dsDNAs. Lastly, it is unknown if Ago limits natural transformation when using exogenous DNA fragments as the substrate. It was recently shown that most SeAgo-associated small DNAs were derived from the proposed genomic replication initiation and termination sites (Olina et al., 2020). Therefore, it is possible that Agos target exogenous DNA when forming DNA intermediates during host cell replication.

CONCLUSIONS AND PERSPECTIVES

The nucleic acid-guided binding and cleavage activities of pAgos are reminiscent of CRISPR-Cas systems. Like CRISPR-Cas systems, pAgos degrade invading DNA into short dsDNA fragments. However, unlike CRISPR-Cas systems, pAgo-mediated defenses do not integrate degraded fragments into host genomes and, thus, no memory of the invading DNA is generated. Therefore, pAgos can be described as the "innate" immune systems of bacteria and archaea (Hegge et al., 2018; Kuzmenko et al., 2019), while the CRISPR-Cas systems can be described as "adaptive" immune systems. The innate and adaptive immune systems interact in mammals, although it is unknown if such interactivity occurs between CRISPR-Cas systems and pAgos. The *ago* genes of *M. piezophila*, *M. kandleri*, and *T. profunda* are present in the same operon as the *cas* gene of the CRISPR-Cas locus (Kaya et al., 2016), although most bacterial and archaeal genomes that encode CRISPR-Cas loci lack Argonaute genes (Makarova et al., 2006). Thus, the potential interrelationships of Ago and CRISPR-Cas systems require further experimental demonstration.

AUTHOR CONTRIBUTIONS

ML, MH, MW, and AC conceived and designed the manuscript. ML, MH, RJ, and SC wrote the paper. MH, DZ, and LZ contributed the figures. ML, MH, and LP supervised the studies and corrected the manuscripts. All the authors have reviewed the manuscript. All the authors who contributed to this work have read the manuscript and approved publication.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 31772772, http://www.nsfc.gov.cn/), Sichuan Science and Technology Program (2020YJ0344), the China Agricultural Research System (CARS-42-17), the Sichuan Veterinary Medicine and Drug Innovation Group of the China Agricultural Research System (SCCXTD-2020-18), and the Integration and Demonstration of Key Technologies for Goose Industrial Chain in Sichuan Province (2018NZ0005).

REFERENCES

Ambur, O. H., Engelstadter, J., Johnsen, P. J., Miller, E. L., and Rozen, D. E. (2016). Steady at the wheel: conservative sex and the benefits of bacterial transformation. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371:20150528. doi: 10.1098/rstb.2015.0528

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

Arber, W. (2014). Horizontal gene transfer among bacteria and its role in biological evolution. *Lifestyles* 4, 217–224. doi: 10.3390/life4020217

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

Balaban, M., Battig, P., Muschiol, S., Tirier, S. M., Wartha, F., Normark, S., et al. (2014). Secretion of a pneumococcal type II secretion system pilus correlates with DNA uptake during transformation. *Proc. Natl. Acad. Sci. U. S. A.* 111, E758–E765. doi: 10.1073/pnas.1313860111

Barrangou, R., and Marraffini, L. A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54, 234–244. doi: 10.1016/j.molcel.2014.03.011

Berge, M. J., Kamgoue, A., Martin, B., Polard, P., Campo, N., and Claverys, J. P. (2013). Midcell recruitment of the DNA uptake and virulence nuclease, EndA, for pneumococcal transformation. *PLoS Pathog.* 9:e1003596. doi: 10.1371/journal.ppat.1003596

Berge, M., Moscoso, M., Prudhomme, M., Martin, B., and Claverys, J. P. (2002). Uptake of transforming DNA in Gram-positive bacteria: a view from *Streptococcus pneumoniae*. *Mol. Microbiol.* 45, 411–421. doi: 10.1046/j.1365-2958.2002.03013.x

Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L. A. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* 12, 177–186. doi: 10.1016/j.chom.2012.06.003

Blokesch, M. (2017). In and out-contribution of natural transformation to the shuffling of large genomic regions. *Curr. Opin. Microbiol.* 38, 22–29. doi: 10.1016/j.mib.2017.04.001

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of *Arabidopsis* controlling leaf development. *EMBO J.* 17, 170–180. doi: 10.1093/emboj/17.1.170

Boonstra, M., Vesel, N., and Kuipers, O. P. (2018). Fluorescently labeled DNA interacts with competence and recombination proteins and is integrated and expressed following natural transformation of *Bacillus subtilis*. *mBio* 9, e01161–e01118. doi: 10.1128/mBio.01161-18

Brodt, A., Lurie-Weinberger, M. N., and Gophna, U. (2011). CRISPR loci reveal networks of gene exchange in archaea. *Biol. Direct* 6:65. doi: 10.1186/1745-6150-6-65

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

Burton, B., and Dubnau, D. (2010). Membrane-associated DNA transport machines. *Cold Spring Harb. Perspect. Biol.* 2:a000406. doi: 10.1101/cshperspect.a000406

Chen, I., and Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* 2, 241–249. doi: 10.1038/nrmicro844

Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439. doi: 10.1126/science.aar6245

Claverys, J. P., Martin, B., and Polard, P. (2009). The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiol. Rev.* 33, 643–656. doi: 10.1111/j.1574-6976.2009.00164.x

Corbinais, C., Mathieu, A., Kortulewski, T., Radicella, J. P., and Marsin, S. (2016). Following transforming DNA in *Helicobacter pylori* from uptake to expression. *Mol. Microbiol.* 101, 1039–1053. doi: 10.1111/mmi.13440

Croucher, N. J., Harris, S. R., Fraser, C., Quail, M. A., Burton, J., van der Linden, M., et al. (2011). Rapid pneumococcal evolution in response to clinical interventions. *Science* 331, 430–434. doi: 10.1126/science.1198545

Dalia, A. B., and Dalia, T. N. (2019). Spatiotemporal analysis of DNA integration during natural transformation reveals a mode of nongenetic inheritance in bacteria. *Cell* 179, 1499.e10–1511.e10. doi: 10.1016/j.cell.2019.11.021

Damke, P. P., Di Guilmi, A. M., Varela, P. F., Velours, C., Marsin, S., Veaute, X., et al. (2019). Identification of the periplasmic DNA receptor for natural transformation of *Helicobacter pylori*. *Nat. Commun.* 10:5357. doi: 10.1038/s41467-019-13352-6

Darmon, E., and Leach, D. R. (2014). Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39. doi: 10.1128/MMBR.00035-13

Draskovic, I., and Dubnau, D. (2005). Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol. Microbiol.* 55, 881–896. doi: 10.1111/j.1365-2958.2004.04430.x

Dubnau, D., and Blokesch, M. (2019). Mechanisms of DNA uptake by naturally competent bacteria. *Annu. Rev. Genet.* 53, 217–237. doi: 10.1146/annurev-genet-112618-043641

Ellison, C. K., Dalia, T. N., Vidal Ceballos, A., Wang, J. C., Biais, N., Brun, Y. V., et al. (2018). Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in *Vibrio cholerae*. *Nat. Microbiol.* 3, 773–780. doi: 10.1038/s41564-018-0174-y

Fischer, S., Maier, L. K., Stoll, B., Brendel, J., Fischer, E., Pfeiffer, F., et al. (2012). An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J. Biol. Chem.* 287, 33351–33363. doi: 10.1074/jbc.M112.377002

Gangel, H., Hepp, C., Muller, S., Oldewurtel, E. R., Aas, F. E., Koomey, M., et al. (2014). Concerted spatio-temporal dynamics of imported DNA and ComE DNA uptake protein during gonococcal transformation. *PLoS Pathog.* 10:e1004043. doi: 10.1371/journal.ppat.1004043

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

Gogarten, J. P., and Townsend, J. P. (2005). Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* 3, 679–687. doi: 10.1038/nrmicro1204

Gophna, U., Kristensen, D. M., Wolf, Y. I., Popa, O., Drevet, C., and Koonin, E. V. (2015). No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *ISME J.* 9, 2021–2027. doi: 10.1038/ismej.2015.20

Hahn, J., Maier, B., Haijema, B. J., Sheetz, M., and Dubnau, D. (2005). Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell* 122, 59–71. doi: 10.1016/j.cell.2005.04.035

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

Hegge, J. W., Swarts, D. C., Chandradoss, S. D., Cui, T. J., Kneppers, J., Jinek, M., et al. (2019). DNA-guided DNA cleavage at moderate temperatures by *Clostridium butyricum* Argonaute. *Nucleic Acids Res.* 47, 5809–5821. doi: 10.1093/nar/gkz306

Hegge, J. W., Swarts, D. C., and van der Oost, J. (2018). Prokaryotic Argonaute proteins: novel genome-editing tools? *Nat. Rev. Microbiol.* 16, 5–11. doi: 10.1038/nrmicro.2017.73

Hepp, C., and Maier, B. (2016). Kinetics of DNA uptake during transformation provide evidence for a translocation ratchet mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12467–12472. doi: 10.1073/pnas.1608110113

Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Hock, J., and Meister, G. (2008). The Argonaute protein family. *Genome Biol.* 9:210. doi: 10.1186/gb-2008-9-2-210

Huang, L., Tian, X., Liu, M., Wang, M., Biville, F., Cheng, A., et al. (2019). DprA is essential for natural competence in *Riemerella anatipestifer* and has a conserved evolutionary mechanism. *Front. Genet.* 10:429. doi: 10.3389/fgene.2019.00429

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

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

Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J. P. (2014). Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* 12, 181–196. doi: 10.1038/nrmicro3199

Johnston, C., Martin, B., Polard, P., and Claverys, J. P. (2013). Postreplication targeting of transformants by bacterial immune systems? *Trends Microbiol.* 21, 516–521. doi: 10.1016/j.tim.2013.08.002

Jorth, P., and Whiteley, M. (2012). An evolutionary link between natural transformation and CRISPR adaptive immunity. *mBio* 3, e00309–e00312. doi: 10.1128/mBio.00309-12

Karnholz, A., Hoefler, C., Odenbreit, S., Fischer, W., Hofreuter, D., and Haas, R. (2006). Functional and topological characterization of novel components of the comB DNA transformation competence system in *Helicobacter pylori*. *J. Bacteriol.* 188, 882–893. doi: 10.1128/JB.188.3.882-893.2006

Kaya, E., Doxzen, K. W., Knoll, K. R., Wilson, R. C., Strutt, S. C., Kranzusch, P. J., et al. (2016). A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4057–4062. doi: 10.1073/pnas.1524385113

Kidane, D., and Graumann, P. L. (2005). Intracellular protein and DNA dynamics in competent *Bacillus subtilis* cells. *Cell* 122, 73–84. doi: 10.1016/j.cell.2005.04.036

Koonin, E. V., and Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374:20180087. doi: 10.1098/rstb.2018.0087

Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008

Kuzmenko, A., Yudin, D., Ryazansky, S., Kulbachinskiy, A., and Aravin, A. A. (2019). Programmable DNA cleavage by Ago nucleases from mesophilic bacteria *Clostridium butyricum* and *Limnothrix rosea*. *Nucleic Acids Res.* 47, 5822–5836. doi: 10.1093/nar/gkz379

Laurenceau, R., Krasteva, P. V., Diallo, A., Ouarti, S., Duchateau, M., Malosse, C., et al. (2015). Conserved *Streptococcus pneumoniae* spirosomes suggest a single type of transformation pilus in competence. *PLoS Pathog.* 11:e1004835. doi: 10.1371/journal.ppat.1004835

Laurenceau, R., Pehau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., et al. (2013). A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog.* 9:e1003473. doi: 10.1371/journal.ppat.1003473

Ma, E., Harrington, L. B., O'Connell, M. R., Zhou, K., and Doudna, J. A. (2015). Single-stranded DNA cleavage by divergent CRISPR-Cas9 enzymes. *Mol. Cell* 60, 398–407. doi: 10.1016/j.molcel.2015.10.030

Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T., and Patel, D. J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434, 666–670. doi: 10.1038/nature03514

Maeusli, M., Lee, B., Miller, S., Reyna, Z., Lu, P., Yan, J., et al. (2020). Horizontal gene transfer of antibiotic resistance from *Acinetobacter baylyi* to *Escherichia coli* on lettuce and subsequent antibiotic resistance transmission to the gut microbiome. *mSphere* 5, e00329–e00320. doi: 10.1128/mSphere.00329-20

Maier, B., Potter, L., So, M., Long, C. D., Seifert, H. S., and Sheetz, M. P. (2002). Single pilus motor forces exceed 100 pN. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16012–16017. doi: 10.1073/pnas.242523299

Makarova, K. S., Gao, L., Zhang, F., and Koonin, E. V. (2019). Unexpected connections between type VI-B CRISPR-Cas systems, bacterial natural competence, ubiquitin signaling network and DNA modification through a distinct family of membrane proteins. *FEMS Microbiol. Lett.* 366:fnz088. doi: 10.1093/femsle/fnz088

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. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577

Makarova, K. S., and Koonin, E. V. (2015). Annotation and classification of CRISPR-Cas systems. *Methods Mol. Biol.* 1311, 47–75. doi: 10.1007/978-1-4939-2687-9_4

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., van der Oost, J., and Koonin, E. V. (2009). Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* 4:29. doi: 10.1186/1745-6150-4-29

Marraffini, L. A., and Sontheimer, E. J. (2010a). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190. doi: 10.1038/nrg2749

Marraffini, L. A., and Sontheimer, E. J. (2010b). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571. doi: 10.1038/nature08703

Matthey, N., and Blokesch, M. (2016). The DNA-uptake process of naturally competent *Vibrio cholerae*. *Trends Microbiol.* 24, 98–110. doi: 10.1016/j.tim.2015.10.008

McCallum, M., Burrows, L. L., and Howell, P. L. (2019). The dynamic structures of the type IV Pilus. *Microbiol. Spectr.* 7:PSIB-0006-2018. doi: 10.1128/microbiolspec.PSIB-0006-2018

Mell, J. C., and Redfield, R. J. (2014). Natural competence and the evolution of DNA uptake specificity. *J. Bacteriol.* 196, 1471–1483. doi: 10.1128/JB.01293-13

Mirouze, N., Ferret, C., Cornilleau, C., and Carballido-Lopez, R. (2018). Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in *Bacillus subtilis*. *Nat. Commun.* 9:5072. doi: 10.1038/s41467-018-07553-8

Miyoshi, T., Ito, K., Murakami, R., and Uchiumi, T. (2016). Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* 7:11846. doi: 10.1038/ncomms11846

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

Mortier-Barriere, I., Velten, M., Dupaigne, P., Mirouze, N., Pietrement, O., McGovern, S., et al. (2007). A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* 130, 824–836. doi: 10.1016/j.cell.2007.07.038

Muschiol, S., Balaban, M., Normark, S., and Henriques-Normark, B. (2015). Uptake of extracellular DNA: competence induced pili in natural transformation of *Streptococcus pneumoniae*. *Bioessays* 37, 426–435. doi: 10.1002/bies.201400125

Nero, T. M., Dalia, T. N., Wang, J. C., Kysela, D. T., Bochman, M. L., and Dalia, A. B. (2018). ComM is a hexameric helicase that promotes branch migration during natural transformation in diverse Gram-negative species. *Nucleic Acids Res.* 46, 6099–6111. doi: 10.1093/nar/gky343

Nivala, J., Shipman, S. L., and Church, G. M. (2018). Spontaneous CRISPR loci generation in vivo by non-canonical spacer integration. *Nat. Microbiol.* 3, 310–318. doi: 10.1038/s41564-017-0097-z

Olina, A., Kuzmenko, A., Ninova, M., Aravin, A. A., Kulbachinskiy, A., and Esyunina, D. (2020). Genome-wide DNA sampling by Ago nuclease from the cyanobacterium *Synechococcus elongatus*. *RNA Biol.* 17, 677–688. doi: 10.1080/15476286.2020.1724716

Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K., and Aravin, A. A. (2013). Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* 51, 594–605. doi: 10.1016/j.molcel.2013.08.014

Overballe-Petersen, S., Harms, K., Orlando, L. A., Mayar, J. V., Rasmussen, S., Dahl, T. W., et al. (2013). Bacterial natural transformation by highly fragmented and damaged DNA. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19860–19865. doi: 10.1073/pnas.1315278110

Peters, L., and Meister, G. (2007). Argonaute proteins: mediators of RNA silencing. *Mol. Cell* 26, 611–623. doi: 10.1016/j.molcel.2007.05.001

Pimentel, Z. T., and Zhang, Y. (2018). Evolution of the natural transformation protein, ComEC, in bacteria. *Front. Microbiol.* 9:2980. doi: 10.3389/fmicb.2018.02980

Provvedi, R., and Dubnau, D. (1999). ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol. Microbiol.* 31, 271–280. doi: 10.1046/j.1365-2958.1999.01170.x

Salzer, R., Kern, T., Joos, F., and Averhoff, B. (2016). The *Thermus thermophilus* comEA/comEC operon is associated with DNA binding and regulation of the DNA translocator and type IV pili. *Environ. Microbiol.* 18, 65–74. doi: 10.1111/1462-2920.12820

Samson, J. E., Magadan, A. H., and Moineau, S. (2015). The CRISPR-Cas immune system and genetic transfers: reaching an equilibrium. *Microbiol. Spectr.* 3, PLAS-0034-2014. doi: 10.1128/microbiolspec.PLAS-0034-2014

Seitz, P., and Blokesch, M. (2013). DNA-uptake machinery of naturally competent Vibrio cholerae. *Proc. Natl. Acad. Sci. U. S. A.* 110, 17987–17992. doi: 10.1073/pnas.1315647110

Seitz, P., Pezeshgi Modarres, H., Borgeaud, S., Bulushev, R. D., Steinbock, L. J., Radenovic, A., et al. (2014). ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable Vibrio cholerae cells. *PLoS Genet.* 10:e1004066. doi: 10.1371/journal.pgen.1004066

Song, J. J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 305, 1434–1437. doi: 10.1126/science.1102514

Stingl, K., Muller, S., Scheidgen-Kleyboldt, G., Clausen, M., and Maier, B. (2010). Composite system mediates two-step DNA uptake into *Helicobacter pylori*. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1184–1189. doi: 10.1073/pnas.0909955107

Swarts, D. C., Hegge, J. W., Hinojo, I., Shiimori, M., Ellis, M. A., Dumrongkulraksa, J., et al. (2015). Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* 43, 5120–5129. doi: 10.1093/nar/gkv415

Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., Snijders, A. P., et al. (2014a). DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* 507, 258–261. doi: 10.1038/nature12971

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V., et al. (2014b). The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* 21, 743–753. doi: 10.1038/nsmb.2879

Swarts, D. C., Szczepaniak, M., Sheng, G., Chandradoss, S. D., Zhu, Y., Timmers, E. M., et al. (2017). Autonomous generation and loading of DNA guides by bacterial Argonaute. *Mol. Cell* 65, 985–998.e986. doi: 10.1016/j.molcel.2017.01.033

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721. doi: 10.1038/nrmicro1234

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D. J. (2008). Structure of the guide-strand-containing argonaute silencing complex. *Nature* 456, 209–213. doi: 10.1038/nature07315

Willkomm, S., Makarova, K. S., and Grohmann, D. (2018). DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.* 42, 376–387. doi: 10.1093/femsre/fuy010

Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., et al. (2005). Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell* 19, 405–419. doi: 10.1016/j.molcel.2005.07.011

Zander, A., Willkomm, S., Ofer, S., van Wolferen, M., Egert, L., Buchmeier, S., et al. (2017). Guide-independent DNA cleavage by archaeal Argonaute from *Methanocaldococcus jannaschii*. *Nat. Microbiol.* 2:17034. doi: 10.1038/nmicrobiol.2017.34

Zhang, X. S., and Blaser, M. J. (2012). Natural transformation of an engineered *Helicobacter pylori* strain deficient in type II restriction endonucleases. *J. Bacteriol.* 194, 3407–3416. doi: 10.1128/JB.00113-12

Zhang, Y., Heidrich, N., Ampattu, B. J., Gunderson, C. W., Seifert, H. S., Schoen, C., et al. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* 50, 488–503. doi: 10.1016/j.molcel.2013.05.001

Zhang, Y., Rajan, R., Seifert, H. S., Mondragon, A., and Sontheimer, E. J. (2015). DNase H activity of *Neisseria meningitidis* Cas9. *Mol. Cell* 60, 242–255. doi: 10.1016/j.molcel.2015.09.020

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.

frontiers in Microbiology

REVIEW
published: 20 January 2021
doi: 10.3389/fmicb.2020.602789

The Cyclic Oligoadenylate Signaling Pathway of Type III CRISPR-Cas Systems

Fengtao Huang and Bin Zhu**

Key Laboratory of Molecular Biophysics, the Ministry of Education, College of Life Science and Technology and Shenzhen College, Huazhong University of Science and Technology, Wuhan, China

Type III CRISPR-Cas systems, which are widespread in both bacteria and archaea, provide immunity against DNA viruses and plasmids in a transcription-dependent manner. Since an unprecedented cyclic oligoadenylate (cOA) signaling pathway was discovered in type III systems in 2017, the cOA signaling has been extensively studied in recent 3 years, which has expanded our understanding of type III systems immune defense and also its counteraction by viruses. In this review, we summarized recent advances in cOA synthesis, cOA-activated effector protein, cOA signaling-mediated immunoprotection, and cOA signaling inhibition, and highlighted the crosstalk between cOA signaling and other cyclic oligonucleotide-mediated immunity discovered very recently.

Keywords: cyclic oligonucleotide, type III systems, CRISPR immune defense, CARF domain proteins, Cas10, CD-NTase

OPEN ACCESS

Edited by:
Aixin Yan,
The University of Hong Kong, Hong Kong

Reviewed by:
Wenyuan Han,
Huazhong Agricultural University, China
Changyi Zhang,
University of Illinois at Urbana-Champaign, United States

****Correspondence:***
Fengtao Huang
huang_fengtao@126.com
Bin Zhu
bin_zhu@hust.edu.cn

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *04 September 2020*
Accepted: *11 November 2020*
Published: *20 January 2021*

Citation:
Huang F and Zhu B (2021) The Cyclic Oligoadenylate Signaling Pathway of Type III CRISPR-Cas Systems. Front. Microbiol. 11:602789. doi: 10.3389/fmicb.2020.602789

INTRODUCTION

CRISPR-Cas systems are known to provide adaptive immunity against viruses and plasmids in prokaryotes. Based on the composition of effector complexes, CRISPR-Cas systems were divided into two classes which could be further subdivided into six types (types I–VI) and multiple subtypes (Makarova et al., 2020b). Class 1 systems (including type I, III, and IV), which have multi-subunit effector complex, are widespread in bacteria and archaea; whereas class 2 (including type II, V, and VI), which contain single-subunit effector complex, are almost completely presented in bacteria (Mohanraju et al., 2016). The effector complexes of type I, II, and V (and possibly IV) target double-stranded DNA (dsDNA), while Type VI system targets RNA (Makarova et al., 2020b). Unlike them, type III effector complex targets both RNA and single-stranded DNA (ssDNA) of the invaders (Tamulaitis et al., 2017). The type III system can be further divided into six subtypes (III A–F), in which Type III-A/D system forms a Csm effector complex composed of five subunits (Csm 1–5) and a single CRISPR RNA (crRNA), while Type III-B/C forms a Cmr effector complex consisting of six subunits (Cmr 1–6) and a crRNA (Makarova et al., 2020b). The effector complexes of type III systems exhibit both target RNA cleavage activity and target RNA-activated ssDNA cleavage activity (Elmore et al., 2016; Estrella et al., 2016; Kazlauskiene et al., 2016). Type III systems provide immunity against invaders depending on the target RNA transcription (Deng et al., 2013; Goldberg et al., 2014). The crRNA-guided Csm/Cmr complexes recognize the complementary target RNA and cleave it into 6 nt nucleotide intervals using the multiple copies of Csm3 or Cmr4 subunit (Hale et al., 2009; Estrella et al., 2016; Kazlauskiene et al., 2016).

Target RNA binding also activates the cyclic oligoadenylate (cOA) synthesis activity of Cas 10 subunit. More details about the transcription-dependent immunity and the structural basis of type III effector complexes and effector proteins had been reviewed elsewhere (Pyenson and Marraffini, 2017; Tamulaitis et al., 2017; Molina et al., 2020). In this review, we systematically discuss the recent advances in cOA signaling pathway of type III systems.

The Cas 10 subunit of type III effector complex and the ancillary ribonuclease Csm6/Csx1 are two important components involved in cOA signaling. Cas 10 contains an N-terminal histidine-aspartate (HD) domain and two Palm domains with a GGDD motif inserted into the second Palm domain (Tamulaitis et al., 2017; **Figure 1A**). The HD domain is responsible for ssDNA cleavage activity, while the Palm domains are homologous to nucleotide polymerases and nucleotide cyclase (Makarova et al., 2002, 2011; Zhu and Ye, 2012), and were hypothesized to synthesize cyclic nucleotides like cyclic di-AMP (Burroughs et al., 2015). However, there was no experimental evidence to verify the domains function for a long time (Koonin and Makarova, 2018). Csm6/Csx1 contains an N-terminal CRISPR-associated Rossman fold (CARF) domain which was predicted to sense nucleotide derivative and a C-terminal higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domain which often functions as ribonuclease (Anantharaman et al., 2013; Makarova et al., 2014; **Figure 1A**). In 2017, two independent studies revealed that the two proteins were involved in a cOA signaling pathway, which had never been found in prokaryotes (Kazlauskiene et al., 2017; Niewoehner et al., 2017). It was found that the Palm domains were responsible for cOA synthesis, and the CARF domain of Csm6/Csx1 can sense the corresponding cOA (Kazlauskiene et al., 2017; Niewoehner et al., 2017; **Figure 1B**). When target RNA is recognized by effector complex, Cas10 subunit can be activated and can generate cOA, which in turn allosterically activates the ribonuclease Csm6/Csx1 through binding the CARF domain, resulting in non-specific RNA degradation (Kazlauskiene et al., 2017; Niewoehner et al., 2017; **Figure 1B**).

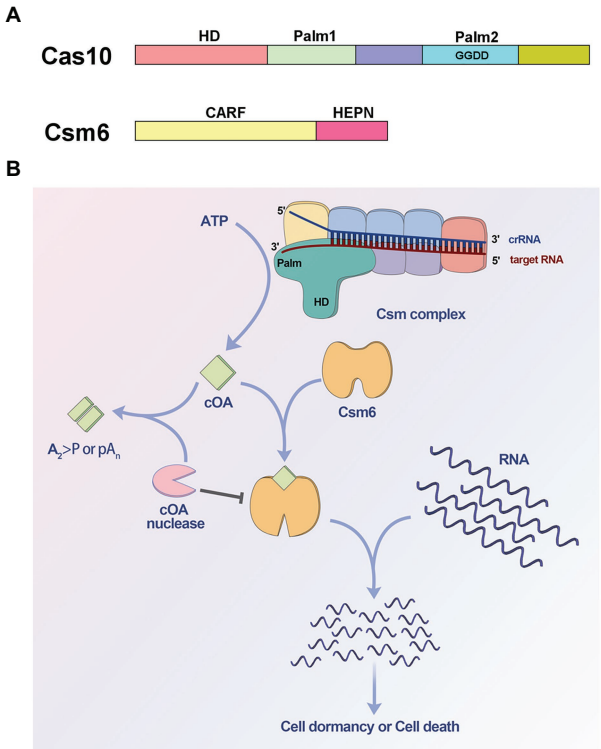


FIGURE 1 | Cyclic oligoadenylate (cOA) signaling-mediated immunity in type III CRISPR-Cas systems. **(A)** Domain organization of Cas10 and Csm6. Cas10 contains an N-terminal histidine-aspartate (HD) domain and two Palm domains (Palm1 and Palm2), and the GGDD motif is inserted into Palm2 domain. Csm6 contains an N-terminal CRISPR-associated Rossman fold (CARF) domain and a C-terminal higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domain. **(B)** Model for cOA signaling pathway of type III systems. The Palm domains of Cas10 subunit are activated and convert ATP into cOA molecules, when the target RNA is recognized by the CRISPR RNA (crRNA)-guided effector complex. The synthesized cOA allosterically activates Csm6 by binding the CARF domain, which subsequently degrades RNA non-specifically, resulting in host cell dormancy or cell death. On the other hand, cOA nucleases including ring nuclease and membrane-associated DHH-DHHA1 family nuclease can degrade cOA molecules to switch off cOA signaling, thereby acting as off-switch for the systems.

Cas10 ACTIVATION-TRIGGERED cOA SYNTHESIS

Cas10 is the largest subunit of type III effector complex and is a signature protein of type III systems (Tamulaitis et al., 2017; Koonin and Makarova, 2018). The cOA synthesis activity of Cas10 is subject to tight spatial and temporal control. The Cas10 subunit is activated and converts ATP into cOA molecules only when target RNA is recognized by the crRNA-guided effector complex, and cOA synthesis will be deactivated abruptly following target RNA cleavage and dissociation from the effector complex (Kazlauskiene et al., 2017; Niewoehner et al., 2017; Rouillon et al., 2018). Unlike type I, II, and V CRISPR-Cas systems which distinguish self from non-self DNA in a protospacer adjacent motif (PAM)-dependent manner, type III systems were proposed to rely on the 5'-handle of crRNA (8 nt) and the 3'-flanking sequence of the target RNA to avoid autoimmunity (Marraffini and Sontheimer, 2010; Kazlauskiene et al., 2016; Tamulaitis et al., 2017). Non-complementarity between crRNA 5'-handle and 3'-flanking sequence of the target RNA is essential for Cas10 activation for cOA synthesis (Kazlauskiene et al., 2017). Previous studies showed type III systems were much tolerant of mismatches in target RNA (Pyenson et al., 2017; Goldberg et al., 2018; Rouillon et al., 2018). To some degree, target RNA binding-mediated Cas10 activation is also tolerant of crRNA-target mismatches, but base pairs in direct contact with Cas10 subunit, such as those adjacent to the 3' end of target RNA, are very stringent (Rouillon et al., 2018; Nasef et al., 2019). Base-pairing of 3'-flanking target RNA sequence to the 5'-handle of crRNA affects activation of both ssDNA cleavage and cOA synthesis (Guo et al., 2019;

Johnson et al., 2019; Foster et al., 2020). Structural studies on Csm complexes show that the interaction between non-complementary 3'-flanking target RNA sequence and Cas10 subunit is crucial to induce a conformational change of Cas10 subunit for activation of its single-stranded DNase (ssDNase) and cOA synthetase activities (Jia et al., 2019c; You et al., 2019). Moreover, recent studies on the Cmr complex show that a unique stalk loop in Cmr3 is critical for avoiding autoimmunity and triggering Cas10 activation (Guo et al., 2019; Sofos et al., 2020). In addition, the length of the crRNA-target duplex also affects Cas10 activation. Twenty five base pairs or longer crRNA-target duplex are required for efficient activation of ssDNA cleavage and cOA generation (You et al., 2019; Sofos et al., 2020).

Since cOA synthesis was identified in Type III-A system of *Streptococcus thermophilus* and *Enterococcus italicus*, respectively (Kazlauskiene et al., 2017; Niewoehner et al., 2017), the Cas10 subunits of effector complexes from various bacteria and *archaea*, which harbor type III-A/B/D systems were verified to generate various cOA molecules (cOA_n, n = 3–6; Han et al., 2018; Rouillon et al., 2018; Grüschow et al., 2019; Nasef et al., 2019; Foster et al., 2020). Notably, the major cOA species produced by effector complex is not always the one that activates the effector ribonuclease in the same system (Kazlauskiene et al., 2017; Rouillon et al., 2019; Smalakyte et al., 2020). That may be because the *in vitro* cOA synthesis could be affected by reaction conditions, thus the synthesized major cOA species may be different from those *in vivo* (Smalakyte et al., 2020). Recently, alternative nucleotide signal molecules were found to be synthesized by GDDEF cyclase, cGAS/DncV-like nucleotidyltransferases (CD-NTases), and ppGpp synthetase homolog (Hallberg et al., 2016; Ahmad et al., 2019; Whiteley et al., 2019), leading us to consider the existence of a subfamily of Cas10-like proteins that can synthesize other kinds of cyclic oligonucleotide molecules. Structure studies on Csm effector complexes bound to substrates (AMPPNP and ATP) have shown that each Palm domain has a conserved serine residue (Ser^{273} and Ser^{549} in the *Streptococcus thermophilus* Csm1), which forms hydrogen bonds with base of ATP and confers specificity for ATP (Jia et al., 2019a; You et al., 2019). Moreover, a biochemical study on Cmr effector complexes also shows that two conserved serine residues in the Palm 1 domain of Cmr2 are important for ATP binding and cOA synthesis, and the study further reveals a cooperative substrate binding mechanism for efficient cOA synthesis (Han et al., 2018). Cas10 with substitutions of the conserved serine residues still retains a certain degree of cOA synthesis activities, yet whether the nucleotide specificity is affected remains unclear (Han et al., 2018; You et al., 2019). It will be interesting to investigate the possibility of other cyclic oligonucleotides synthesis from uncharacterized type III effector complexes.

cOA-ACTIVATED EFFECTOR PROTEINS

The effector proteins Csm6 and Csx1, representatives of CARF family proteins, can be activated by either cOA_4 or cOA_6, depending on their preferences (Shah et al., 2019). Very recently, the crystal structures of complexes of Csm6/Csx1 with cOA molecules have been determined (Jia et al., 2019b; Molina et al., 2019; Garcia-Doval et al., 2020). Studies on these structures reveal that one cOA binds to each CARF domains of the symmetrical homodimer of Csm6/Csx1, resulting in conformational change of Csm6/Csx1 and HEPN domain activation (Jia et al., 2019b; Molina et al., 2019; Garcia-Doval et al., 2020). Furthermore, the CARF domain of Csm6 can autoregulate its RNase activity through degrading its cOA activators (Athukoralage et al., 2019; Jia et al., 2019b; Garcia-Doval et al., 2020). However, the CARF domain of Csx1 cannot cleave its cOA activator, suggesting that the degradation of cOA by CARF domains is not a general mechanism for CARF family proteins (Molina et al., 2019).

Since CARF domain is responsible for sensing cOA, other CARF domain-containing proteins may also serve as the effector proteins. Bioinformatics analysis shows that the CARF domain is also fused to various other domains in type III systems, implying cOA signaling may provide immunity through activating various CARF domain proteins not just Csm6/Csx1 (Makarova et al., 2014, 2020a; Koonin and Makarova, 2018; Shah et al., 2019). For examples, CARF domain is fused to other RNase domains such as ribosome-dependent endoribonuclease RelE and PIN, and DNase domains of restriction endonuclease (REase) and HD nuclease, suggesting that RNA and even DNA can be degraded by such cOA-activated CARF domain proteins (Makarova et al., 2014, 2020a; Koonin and Makarova, 2018). Additionally, CARF domains are also fused with domains such as helix-turn-helix (HTH), AAA+ ATPase, or adenosine deaminase, suggesting that RNA transcription can also be regulated by such cOA-activated CARF domain proteins (Makarova et al., 2014, 2020a; Koonin and Makarova, 2018). Very recently, an effector protein containing two CARF domains and one DNA nuclease-like domain (named Can1) and another effector protein containing a Csx1 protein fused to a ring nuclease CRISPR-associated ring nuclease 2 (Crn2) domain (named Csx1-Crn2) are characterized (McMahon et al., 2020; Samolygo et al., 2020). Unlike Csm6/Csx1, Can1 is a monomeric enzyme with DNA nuclease activity (McMahon et al., 2020), while Csx1-Crn2 degrades cOA_4 by the Crn2 domain to limit its cOA_4-activated ribonuclease activity (Samolygo et al., 2020). These results demonstrated the diversity of CARF domain-containing effectors.

It has been known that activated Cas10 subunit produces cOAs ranging from cOA_3 to cOA_6 (Kazlauskiene et al., 2017). However, it is unlikely that cOA_3 or cOA_5 can activate CARF domain proteins like Csm6/Csx1 which assembles as homodimer with 2-fold symmetry, because the two cOAs lack symmetry to fit the dimer interface of CARF domains (Rouillon et al., 2018). Thus, it was questioned why type III effector complex generates cOA_3 and cOA_5, which are even the predominant products (Kazlauskiene et al., 2017; Smalakyte et al., 2020); and whether there are any other kinds of effector proteins presented in type III systems.

Very recently, it is found that a novel CD-NTase produces cOA_3, which in turn activates its effector endonuclease NucC to degrade DNA non-specifically to provide immunity against bacteriophage (Lau et al., 2020; Ye et al., 2020). Interestingly, NucC homologs as accessory proteins are also encoded within type III CRISPR/Cas systems and can be strongly activated by cOA_3, indicating the existence of effector proteins without CARF domains in type III systems (Lau et al., 2020; Malone et al., 2020). Indeed, many other kinds of accessory proteins have been identified in type III systems (Shah et al., 2019). New cOA-activated effector proteins may still exist and remain to be identified, especially in some type III systems that contain Cas10 but lack any CARF domain proteins (Koonin and Makarova, 2018).

IMMUNOPROTECTION CONFERRED BY cOA SIGNALING

The effector protein Csm6 has been shown to be essential for type III-A CRISPR-Cas systems against phage and plasmid even before the cOA signaling was discovered in 2017 (Hatoum-Aslan et al., 2014; Jiang et al., 2016). Anti-phage activity of Csm6 was demonstrated to be dependent on Cas10 activation and cOA synthesis *in vivo* at the time of cOA signaling discovery (Niewoehner et al., 2017). Since Csm6/Csx1 cleaves RNA with a preference for only one or two nucleotides (Kazlauskiene et al., 2017; Foster et al., 2019; Jia et al., 2019b; Molina et al., 2019), it is largely sequence non-specific, and degrades RNA of host and invader indiscriminately. Thus, the activated effector Csm6/Csx1 is deleterious to the host, and it was proposed that cOA signaling confers host defense through inducing cell dormancy to arrest infection or inducing programmed host cell death to abort infection (Kazlauskiene et al., 2017). Indeed, it has been observed that Csm6 activation resulted in degradation of both host and plasmid transcripts, and induced growth arrest of the host which was critical for plasmid clearance (Rostøl and Marraffini, 2019). Recently, it was found that a kind of jumbo phages form nucleus-like structures during infection to protect their DNA from DNA-targeting nucleases (Chaikeeratisak et al., 2017; Mendoza et al., 2020). However, a type III system can provide robust immunity against such nucleus-forming jumbo phage (Malone et al., 2020). In this case, the cOA signaling is essential for the type III system against the jumbo phage (Malone et al., 2020). Interestingly, the effector protein involved in the cOA signaling is a NucC-like DNA nuclease but not the ribonuclease (Malone et al., 2020), which is not capable of cleaving the jumbo phage DNA in principle. Thus, it is likely that the cOA signaling confers defense by inducing host dormancy or abortive infection through non-specifically degrading the host genome.

It is worth noting that Csm6/Csx1 activation is crucial for efficient immunity against virus when targets are late-expressed viral genes but not the early-expressed genes (Jiang et al., 2016; Bhoobalan-Chitty et al., 2019). It was suggested that the Cas10 ssDNase is sufficient to clear the invaders when targets are early-expressed genes (Jiang et al., 2016; Bhoobalan-Chitty et al., 2019). In this case, it is not necessary for the host to activate cOA signaling pathway which might also be toxic to the host. Indeed, it was shown that targeting the late-expressed viral gene exhibits a relatively stronger antiviral immunity than targeting the early-expressed viral gene when the Cas10 ssDNase is inactivated, indicating that cOA signaling-mediated immunity may be stronger in targeting the late-expressed viral gene than the early-expressed gene (Bhoobalan-Chitty et al., 2019). Recent studies have shown that the Palm domain of Cas10 can be strongly activated even when the target genes are transcribed at very low levels (Rostøl and Marraffini, 2019; Athukoralage et al., 2020b), so it is unlikely that transcripts from early-expressed genes cannot activate Cas10. One possibility is that the activity of Cas10 Palm domain is inhibited in the early infection stage by some unknown mechanism in host cells, and this inhibition is released in the late infection stage for the activation of Cas10 Palm domain. Such hypothesis could be supported by findings that cellular nucleotides such as dATP, AMP, and ADP can also bind to the adenosine binding sites of Cas10 and affect cOA synthesis (Kazlauskiene et al., 2017). The level of these nucleotides may be decreased during viral replication, promoting activation of cOA synthesis.

Histidine-aspartate domain of Cas10, which is responsible for non-specific ssDNA degradation, is also involved in immunity against viruses and plasmids. It was reported that cOA signaling should be coupled with Cas10 ssDNase activity for efficient clearance of invader genomes (Jiang et al., 2016; Rostøl and Marraffini, 2019; Varble and Marraffini, 2019). However, in some studies, inactivation of HD domain of Cas10 has little effect on immunity against invaders, suggesting that without the assistance of Cas10 ssDNase activity, the type III effector complex with cOA signaling can still provide sufficient immunoprotection (Foster et al., 2019; Liu et al., 2019; Malone et al., 2020). Cas10 ssDNase was previously proposed to be involved in ssDNA cleavage at the transcription bubble, but a recent study argued this mechanism, leaving the real role of the Cas10 ssDNase unclear (Liu et al., 2019). Thus, how cOA signaling cooperates with the Cas10 ssDNase for immune defense remains to be investigated.

cOA SIGNALING INHIBITION

Due to its toxicity to the host, extant cOA should be removed after clearance of the invaders to enable host cells to return to normal growth state (**Figure 1B**). The first cOA nuclease, also named CRISPR-associated ring nuclease 1 (Crn1), was identified from crenarchaeote *Sulfolobus solfataricus* (Athukoralage et al., 2018). Crn1 is a CARF domain-containing protein that forms a homodimer, and specifically cleaves cOA_4 into linear di-adenylate products to switch off the cOA_4-activated effector proteins (Athukoralage et al., 2018). Interestingly, the

CARF domain of effector protein Csm6 is also found to be capable of degrading cOA, thereby functioning as self-limiting ribonucleases (Athukoralage et al., 2019; Jia et al., 2019b; Garcia-Doval et al., 2020). Notably, very recently, it was found the HEPN domain of Csm6 can also degrade cOA to self-regulate its RNase activity (Smalakyte et al., 2020). Moreover, recent studies report that the widespread CRISPR associated protein Csx3 is a novel ring nuclease, named Crn3 (CRISPR associated ring nuclease 3; Athukoralage et al., 2020a; Brown et al., 2020). Interestingly, an unusual cooperative catalytic mechanism was found in which an active site of Csx3 tetramer is formed by two dimers sandwiching a cOA_4 substrate (Athukoralage et al., 2020a). In addition, a metal-dependent and membrane-associated DHH-DHHA1 family nuclease (MAD) from *Sulfolobus islandicus* has recently been identified as a novel cOA-degrading enzyme (Zhao et al., 2020). MAD can accelerate the clearance of high-level cOA and may cooperate with cellular ring nuclease to remove cOA (Zhao et al., 2020).

Since cOA signaling promotes strong antiviral immunity, conversely, virus can utilize different strategies to restrict cOA signaling for immune evasion. Obviously, cOA degradation is a simple and efficient way for viruses to evade immune response. Indeed, a new family of viral anti-CRISPR (Acr) protein, AcrIII-1, was recently identified as a ring nuclease that specifically degrades cOA_4, suggesting that it functions as Acr protein against cOA_4-triggered type III CRISPR-Cas immunity (Athukoralage et al., 2020c). AcrIII-1 has a higher activity for cOA_4 degradation than Crn1 and Crn3, and is unrelated to the CARF family proteins (Athukoralage et al., 2020a,c). AcrIII-1 homologs are widespread in various prokaryotes, where AcrIII-1 homologs may function as host-encoded ring nuclease like Crn1 and Crn3, thus named Crn2 (Athukoralage et al., 2020c). Very recently, a novel type III CRISPR-Cas inhibitor AcrIIIB1, encoded by *Sulfolobus* virus, has been identified to inhibit type III-B system immunity by binding to its effector complex to affect cOA signaling, demonstrating another strategy developed by virus to evade cOA signaling-mediated immunity (Bhoobalan-Chitty et al., 2019).

Currently, all of the characterized cOA nucleases (Crn1–3 and MAD) specifically degrade cOA_4. However, given that Cas10 synthesizes cOAs ranging from cOA_3 to cOA_6, it is rational to predict the existence of other cOA-specific nucleases in prokaryotes and viruses. A cOA_6-activated Csm6 can cleave cOA_6 by its CARF domain (Garcia-Doval et al., 2020), indicating there may be presence of other CARF domain proteins such as Crn1 homologs that can specifically cleave cOA_6. Moreover, MAD which is distinct from ring nucleases has a board substrate spectrum including cyclic di-nucleotides and ssRNA, implying MAD could degrade various cOAs (Zhao et al., 2020). Additionally, AcrIIIB1 utilizes a special strategy to inhibit cOA signaling, but its homologous proteins are only found in a few archaeal viruses (Bhoobalan-Chitty et al., 2019). The inhibitory activity of AcrIIIB1 also inspires us to explore other Acr proteins in bacteriophages and more strategies for evasion of cOA signaling such as inhibiting the effector protein activities.

CONCLUDING REMARKS

In recent studies, a large family of CD-NTases have been found to produce a wide variety of cyclic di- and trinucleotides including 3'3' cyclic UMP-AMP, 3'3'3' cyclic AMP-AMP-GMP, and 2'3'3' cyclic AMP-AMP-AMP, which had never been reported previously (Whiteley et al., 2019; Lowey et al., 2020). These cyclic di- and trinucleotides can activate the downstream effector proteins, such as patatin-like phospholipases, DNA endonucleases, proteases, and pore-forming transmembrane proteins, to mediate anti-phage immunity by abortive infection (Cohen et al., 2019; Lau et al., 2020; Lowey et al., 2020; Ye et al., 2020), which is similar to the action of effector proteins of cOA signaling in type III systems. This newly discovered anti-bacteriophage defense system is termed cyclic oligonucleotide-based anti-phage signaling system (CBASS; Cohen et al., 2019), which is widespread and diverse in bacteria and crosstalks with cOA signaling. For example, cOA_3-activated DNA endonuclease in CBASS is also present in some type III systems and is considered to be an important effector protein of cOA signaling for immunity against phage (Lau et al., 2020; Malone et al., 2020). Very recently, it was found that the major CBASS-associated protein effectors contain a SAVED domain, which is a fusion of two CARF-like domains, but recognize diverse asymmetric cyclic oligonucleotide signals such as 3'3'3'cyclic AMP-AMP-GMP and 3'3'3' cyclic AMP-AMP-AMP (cOA_3) which are synthesized by CD-NTases (Lowey et al., 2020). Interestingly, Bioinformatics analysis showed that CD-NTases and SAVED domains fused to protein partners such as Lon protease and pore-forming transmembrane protein, which are occasionally incorporated into type III CRISPR loci (Burroughs et al., 2015; Lowey et al., 2020), which may increase the complexity of the regulation of the cOA signaling and enhance cOA signaling-mediated immunity.

Cyclic oligonucleotide molecules discovered in the past 3 years have greatly expanded our understanding on nucleotide signaling molecules in prokaryotes for decades. Due to the diversity of the effector proteins in cOA signaling and crosstalk between cOA signaling and CBASS, cOA signaling is far from fully elucidated. Future studies of cOA regulation will further expand our understanding of the role of cOA signaling and give us new insights into the cyclic oligonucleotides involved in antiviral defense systems.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31900032, 31670175, and 31870165) and Shenzhen Science and Technology Innovation Fund (JCYJ20170413115637100).

REFERENCES

Ahmad, S., Wang, B., Walker, M. D., Tran, H. R., Stogios, P. J., Savchenko, A., et al. (2019). An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. *Nature* 575, 674–678. doi: 10.1038/s41586-019-1735-9

Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct* 8:15. doi: 10.1186/1745-6150-8-15

Athukoralage, J. S., Graham, S., Grüschow, S., Rouillon, C., and White, M. F. (2019). A type III CRISPR ancillary ribonuclease degrades its cyclic oligoadenylate activator. *J. Mol. Biol.* 431, 2894–2899. doi: 10.1016/j.jmb.2019.04.041

Athukoralage, J. S., Graham, S., Rouillon, C., Grüschow, S., Czekster, C. M., and White, M. F. (2020b). The dynamic interplay of host and viral enzymes in type III CRISPR-mediated cyclic nucleotide signalling. *Elife* 9:e55852. doi: 10.7554/eLife.55852

Athukoralage, J. S., McMahon, S. A., Zhang, C., Grüschow, S., Graham, S., Krupovic, M., et al. (2020c). An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. *Nature* 577, 572–575. doi: 10.1038/s41586-019-1909-5

Athukoralage, J. S., McQuarrie, S., Grüschow, S., Graham, S., Gloster, T. M., and White, M. F. (2020a). Tetramerisation of the CRISPR ring nuclease Crn3/Csx3 facilitates cyclic oligoadenylate cleavage. *Elife* 9:e57627. doi: 10.7554/eLife.57627

Athukoralage, J. S., Rouillon, C., Graham, S., Grüschow, S., and White, M. F. (2018). Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* 562, 277–280. doi: 10.1038/s41586-018-0557-5

Bhoobalan-Chitty, Y., Johansen, T. B., Di Cianni, N., and Peng, X. (2019). Inhibition of type III CRISPR-Cas immunity by an archaeal virus-encoded anti-CRISPR protein. *Cell* 179, 448–458. doi: 10.1016/j.cell.2019.09.003

Brown, S., Gauvin, C. C., Charbonneau, A. A., Burman, N., and Lawrence, C. M. (2020). Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type-III CRISPR-Cas that degrades the second messenger cA4. *J. Biol. Chem.* 295, 14963–14972. doi: 10.1074/jbc.RA120.014099

Burroughs, A. M., Zhang, D., Schäffer, D. E., Iyer, L. M., and Aravind, L. (2015). Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* 43, 10633–10654. doi: 10.1093/nar/gkv1267

Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K., et al. (2017). Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355, 194–197. doi: 10.1126/science.aal2130

Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., et al. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* 574, 691–695. doi: 10.1038/s41586-019-1605-5

Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013). A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol. Microbiol.* 87, 1088–1099. doi: 10.1111/mmi.12152

Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR-Cas system. *Genes Dev.* 30, 447–459. doi: 10.1101/gad.272153.115

Estrella, M. A., Kuo, F. T., and Bailey, S. (2016). RNA-activated DNA cleavage by the type III-B CRISPR-Cas effector complex. *Genes Dev.* 30, 460–470. doi: 10.1101/gad.273722.115

Foster, K., Grüschow, S., Bailey, S., White, M. F., and Terns, M. P. (2020). Regulation of the RNA and DNA nuclease activities required for Pyrococcus furiosus type III-B CRISPR-Cas immunity. *Nucleic Acids Res.* 48, 4418–4434. doi: 10.1093/nar/gkaa176

Foster, K., Kalter, J., Woodside, W., Terns, R. M., and Terns, M. P. (2019). The ribonuclease activity of Csm6 is required for anti-plasmid immunity by type III-A CRISPR-Cas systems. *RNA Biol.* 16, 449–460. doi: 10.1080/15476286.2018.1493334

Garcia-Doval, C., Schwede, F., Berk, C., Rostøl, J. T., Niewoehner, O., Tejero, O., et al. (2020). Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. *Nat. Commun.* 11:1596. doi: 10.1038/s41467-020-15334-5

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

Goldberg, G. W., McMillan, E. A., Varble, A., Modell, J. W., Samai, P., Jiang, W., et al. (2018). Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces the fitness of lysogenic hosts. *Nat. Commun.* 9:61. doi: 10.1038/s41467-017-02557-2

Grüschow, S., Athukoralage, J. S., Graham, S., Hoogeboom, T., and White, M. F. (2019). Cyclic oligoadenylate signalling mediates *Mycobacterium tuberculosis* CRISPR defence. *Nucleic Acids Res.* 47, 9259–9270. doi: 10.1093/nar/gkz676

Guo, T., Zheng, F., Zeng, Z., Yang, Y., Li, Q., She, Q., et al. (2019). Cmr3 regulates the suppression on cyclic oligoadenylate synthesis by tag complementarity in a type III-B CRISPR-Cas system. *RNA Biol.* 16, 1513–1520. doi: 10.1080/15476286.2019.1642725

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

Hallberg, Z. F., Wang, X. C., Wright, T. A., Nan, B., Ad, O., Yeo, J., et al. (2016). Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3', 3'-cGAMP). *Proc. Natl. Acad. Sci. U. S. A.* 113, 1790–1795. doi: 10.1073/pnas.1515287113

Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., Peng-Lundgren, L., et al. (2018). A type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. *Nucleic Acids Res.* 46, 10319–10330. doi: 10.1093/nar/gky844

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

Jia, N., Jones, R., Sukenick, G., and Patel, D. J. (2019a). Second messenger cA(4) formation within the composite Csm1 palm pocket of type III-A CRISPR-Cas Csm complex and its release path. *Mol. Cell* 75, 933–943. doi: 10.1016/j.molcel.2019.06.013

Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D. J. (2019b). CRISPR-Cas III-A Csm6 CARF domain is a ring nuclease triggering stepwise cA(4) cleavage with ApA>p formation terminating RNase activity. *Mol. Cell* 75, 944–956. doi: 10.1016/j.molcel.2019.06.014

Jia, N., Mo, C. Y., Wang, C., Eng, E. T., Marraffini, L. A., and Patel, D. J. (2019c). Type III-A CRISPR-Cas Csm complexes: assembly, periodic RNA cleavage, DNase activity regulation, and autoimmunity. *Mol. Cell* 73, 264–277.e5. doi: 10.1016/j.molcel.2018.11.007

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

Johnson, K., Learn, B. A., Estrella, M. A., and Bailey, S. (2019). Target sequence requirements of a type III-B CRISPR-Cas immune system. *J. Biol. Chem.* 294, 10290–10299. doi: 10.1074/jbc.RA119.008728

Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* 357, 605–609. doi: 10.1126/science.aao0100

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

Koonin, E. V., and Makarova, K. S. (2018). Discovery of oligonucleotide signaling mediated by CRISPR-associated polymerases solves two puzzles but leaves an enigma. *ACS Chem. Biol.* 13, 309–312. doi: 10.1021/acschembio.7b00713

Lau, R. K., Ye, Q., Birkholz, E. A., Berg, K. R., Patel, L., Mathews, I. T., et al. (2020). Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity. *Mol. Cell* 77, 723–733. doi: 10.1016/j.molcel.2019.12.010

Liu, T. Y., Liu, J. J., Aditham, A. J., Nogales, E., and Doudna, J. A. (2019). Target preference of type III-A CRISPR-Cas complexes at the transcription bubble. *Nat. Commun.* 10:3001. doi: 10.1038/s41467-019-10780-2

Lowey, B., Whiteley, A. T., Keszei, A. F. A., Morehouse, B. R., Mathews, I. T., Antine, S. P., et al. (2020). CBASS immunity uses CARF-related effectors to sense 3'-5'- and 2'-5'-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* 182, 38.e17–49.e17. doi: 10.1016/j.cell.2020.05.019

Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014). CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front. Genet.* 5:102. doi: 10.3389/fgene.2014.00102

Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B., and Koonin, E. V. (2002). A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* 30, 482–496. doi: 10.1093/nar/30.2.482

Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011). 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., Timinskas, A., Wolf, Y. I., Gussow, A. B., Siksnys, V., Venclovas, Č., et al. (2020a). Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antivirus defense. *Nucleic Acids Res.* 48, 8828–8847. doi: 10.1093/nar/gkaa635

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020b). 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

Malone, L. M., Warring, S. L., Jackson, S. A., Warnecke, C., Gardner, P. P., Gumy, L. F., et al. (2020). A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat. Microbiol.* 5, 48–55. doi: 10.1038/s41564-019-0612-5

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

McMahon, S. A., Zhu, W., Graham, S., Rambo, R., White, M. F., and Gloster, T. M. (2020). Structure and mechanism of a type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. *Nat. Commun.* 11:500. doi: 10.1038/s41467-019-14222-x

Mendoza, S. D., Nieweglowska, E. S., Govindarajan, S., Leon, L. M., Berry, J. D., Tiwari, A., et al. (2020). A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* 577, 244–248. doi: 10.1038/s41586-019-1786-y

Mohanraju, P., Makarova, K. S., Zetsche, B., Zhang, F., Koonin, E. V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353:aad5147. doi: 10.1126/science.aad5147

Molina, R., Sofos, N., and Montoya, G. (2020). Structural basis of CRISPR-Cas type III prokaryotic defence systems. *Curr. Opin. Struct. Biol.* 65, 119–129. doi: 10.1016/j.sbi.2020.06.010

Molina, R., Stella, S., Feng, M., Sofos, N., Jauniskis, V., Pozdnyakova, I., et al. (2019). Structure of Csx1-cOA$_4$ complex reveals the basis of RNA decay in type III-B CRISPR-Cas. *Nat. Commun.* 10:4302. doi: 10.1038/s41467-019-12244-z

Nasef, M., Muffly, M. C., Beckman, A. B., Rowe, S. J., Walker, F. C., Hatoum-Aslan, A., et al. (2019). Regulation of cyclic oligoadenylate synthesis by the *Staphylococcus epidermidis* Cas10-Csm complex. *RNA* 25, 948–962. doi: 10.1261/rna.070417.119

Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., Bigler, L., et al. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548, 543–548. doi: 10.1038/nature23467

Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O., and Marraffini, L. A. (2017). Broad targeting specificity during bacterial type III CRISPR-Cas immunity constrains viral escape. *Cell Host Microbe* 22, 343.e3–353.e3. doi: 10.1016/j.chom.2017.07.016

Pyenson, N. C., and Marraffini, L. A. (2017). Type III CRISPR-Cas systems: when DNA cleavage just isn't enough. *Curr. Opin. Microbiol.* 37, 150–154. doi: 10.1016/j.mib.2017.08.003

Rostøl, J. T., and Marraffini, L. A. (2019). Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat. Microbiol.* 4, 656–662. doi: 10.1038/s41564-018-0353-x

Rouillon, C., Athukoralage, J. S., Graham, S., Grüschow, S., and White, M. F. (2018). Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* 7:e36734. doi: 10.7554/eLife.36734

Rouillon, C., Athukoralage, J. S., Graham, S., Grüschow, S., and White, M. F. (2019). Investigation of the cyclic oligoadenylate signaling pathway of type III CRISPR systems. *Methods Enzymol.* 616, 191–218. doi: 10.1016/bs.mie.2018.10.020

Samolygo, A., Athukoralage, J. S., Graham, S., and White, M. F. (2020). Fuse to defuse: a self-limiting ribonuclease-ring nuclease fusion for type III CRISPR defence. *Nucleic Acids Res.* 48, 6149–6156. doi: 10.1093/nar/gkaa298

Shah, S. A., Alkhnbashi, O. S., Behler, J., Han, W., She, Q., Hess, W. R., et al. (2019). Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families. *RNA Biol.* 16, 530–542. doi: 10.1080/15476286.2018.1483685

Smalakyte, D., Kazlauskiene, M., Havelund, J. F., Rukšėnaitė, A., Rimaite, A., Tamulaitiene, G., et al. (2020). Type III-A CRISPR-associated protein Csm6 degrades cyclic hexa-adenylate activator using both CARF and HEPN domains. *Nucleic Acids Res.* 48, 9204–9217. doi: 10.1093/nar/gkaa634

Sofos, N., Feng, M., Stella, S., Pape, T., Fuglsang, A., Lin, J., et al. (2020). Structures of the Cmr-β complex reveal the regulation of the immunity mechanism of type III-B CRISPR-Cas. *Mol. Cell* 79, 741.e7–757.e7. doi: 10.1016/j.molcel.2020.07.008

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

Varble, A., and Marraffini, L. A. (2019). Three new Cs for CRISPR: collateral, communicate, cooperate. *Trends Genet.* 35, 446–456. doi: 10.1016/j.tig.2019.03.009

Whiteley, A. T., Eaglesham, J. B., de Oliveira Mann, C. C., Morehouse, B. R., Lowey, B., Nieminen, E. A., et al. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* 567, 194–199. doi: 10.1038/s41586-019-0953-5

Ye, Q., Lau, R. K., Mathews, I. T., Birkholz, E. A., Watrous, J. D., Azimi, C. S., et al. (2020). HORMA domain proteins and a trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity. *Mol. Cell* 77, 709.e7–722.e7. doi: 10.1016/j.molcel.2019.12.009

You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., et al. (2019). Structure studies of the CRISPR-Csm complex reveal mechanism of co-transcriptional interference. *Cell* 176, 239–253. doi: 10.1016/j.cell.2018.10.052

Zhao, R., Yang, Y., Zheng, F., Zeng, Z., Feng, W., Jin, X., et al. (2020). A membrane-associated DHH-DHHA1 nuclease degrades type III CRISPR second messenger. *Cell Rep.* 32:108133. doi: 10.1016/j.celrep.2020.108133

Zhu, X., and Ye, K. (2012). Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. *FEBS Lett.* 586, 939–945. doi: 10.1016/j.febslet.2012.02.036

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 March 2021
doi: 10.3389/fmicb.2021.657404

The CRISPR-Cas System Is Involved in OmpR Genetic Regulation for Outer Membrane Protein Synthesis in *Salmonella* Typhi

*Liliana Medina-Aparicio[1], Sarahí Rodriguez-Gutierrez[1], Javier E. Rebollar-Flores[1], Ángel G. Martínez-Batallar[2], Blanca D. Mendoza-Mejía[1], Eira D. Aguirre-Partida[1], Alejandra Vázquez[1], Sergio Encarnación[2], Edmundo Calva[1] and Ismael Hernández-Lucas[1]**

[1]*Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Mexico,* [2]*Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Mexico*

OPEN ACCESS

Edited by:
Graciela Castro Escarpulli, Instituto Politécnico Nacional de México (IPN), Mexico

Reviewed by:
Josep Casadesús, Sevilla University, Spain
Rolf Dieter Joerger, University of Delaware, United States

***Correspondence:**
Ismael Hernández-Lucas
ismael.hernandez@ibt.unam.mx; ismaelh@ibt.unam.mx

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *22 January 2021*
Accepted: *10 March 2021*
Published: *29 March 2021*

Citation:
Medina-Aparicio L, Rodriguez-Gutierrez S, Rebollar-Flores JE, Martínez-Batallar AG, Mendoza-Mejía BD, Aguirre-Partida ED, Vázquez A, Encarnación S, Calva E and Hernández-Lucas I (2021) The CRISPR-Cas System Is Involved in OmpR Genetic Regulation for Outer Membrane Protein Synthesis in Salmonella Typhi. Front. Microbiol. 12:657404. doi: 10.3389/fmicb.2021.657404

The CRISPR-Cas cluster is found in many prokaryotic genomes including those of the Enterobacteriaceae family. *Salmonella enterica* serovar Typhi (*S*. Typhi) harbors a Type I-E CRISPR-Cas locus composed of *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2*, and a CRISPR1 array. In this work, it was determined that, in the absence of *cas5* or *cas2*, the amount of the OmpC porin decreased substantially, whereas in individual *cse2*, *cas6e*, *cas1*, or *cas3* null mutants, the OmpF porin was not observed in an electrophoretic profile of outer membrane proteins. Furthermore, the LysR-type transcriptional regulator LeuO was unable to positively regulate the expression of the quiescent OmpS2 porin, in individual *S*. Typhi *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* mutants. Remarkably, the expression of the master porin regulator OmpR was dependent on the Cse2, Cas5, Cas6e, Cas1, Cas2, and Cas3 proteins. Therefore, the data suggest that the CRISPR-Cas system acts hierarchically on OmpR to control the synthesis of outer membrane proteins in *S*. Typhi.

Keywords: CRISPR-Cas, porin regulation, *Salmonella* Typhi, OmpR, outer membrane proteins

INTRODUCTION

Microorganisms are constantly exposed to multiple viral infections and have developed many strategies to survive phage attack and invasion by foreign DNA. One such strategy is the CRISPR-Cas bacterial immunological system (Barrangou et al., 2007). This system is classified according to the presence of signature Cas proteins (Makarova et al., 2011, 2015). The hallmark of the CRISPR-Cas Type I system is the presence of the endonuclease Cas3. This protein is involved in cleavage of exogenous target nucleic acids (Sinkunas et al., 2011; Westra et al., 2012). The Type II system requires Cas9 and a trans-activating CRISPR RNA (tracrRNA) for DNA recognition and degradation (Deltcheva et al., 2011). The Type III system uses the RAMP proteins and Cas10 nuclease to silence the invader (Samai et al., 2015; Elmore et al., 2016).

In the Enterobacteriaceae family, the Type I CRISPR-Cas is the predominant system. The analysis of 228 enterobacterial genomes, corresponding to 38 genera, showed that 55% present, at least, one Type I CRISPR-Cas system (Medina-Aparicio et al., 2018). In the *Salmonella* genus, two CRISPR arrays (CRISPR1 and CRISPR2) have been identified, and only CRISPR1 is associated

with a Type I-E set of *cas* genes (Touchon and Rocha, 2010). In 35 of 38 *Salmonella* genomes analyzed so far, the Type I-E CRISPR-Cas system was present, whereas *S. enterica* serovars Pullorum S06004, Javiana and Paratyphi B did not have any *cas* genes (Medina-Aparicio et al., 2018).

Salmonella Typhi IMSS-1 harbors a Type I-E CRISPR-Cas cluster composed of *cas3*, *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*, an 84-bp leader sequence, seven 29-bp repeats and six 32-bp spacers with no homologous sequences reported in the DDBJ data bank (Medina-Aparicio et al., 2011). This locus contains five transcriptional units, two of them are the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR (*cas*-CRISPR operon) and s*cse2* (sense *cse2* RNA), are transcribed from the sense strand, whereas as*cse2-1* (antisense RNA of *cse2* to *cse1*) and as*cas2-1* (antisense RNA of *cas2* to *cas1*) are present on the antisense strand (Medina-Aparicio et al., 2017). Additionally, the *S.* Typhi *cas3* gene is transcribed as an independent unit divergent from the *cas*-CRISPR operon (**Figure 1**). The transcription of the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR polycistronic mRNA is induced by LeuO and negatively regulated by H-NS and Lrp (Hernández-Lucas et al., 2008; Medina-Aparicio et al., 2011). The role of H-NS in silencing the expression of the *cas3* and as*cse2-1* transcriptional units has also been demonstrated. The transcriptional activities of the five transcriptional units present in the *S.* Typhi CRISPR-Cas locus are induced by basic pH (Medina-Aparicio et al., 2017).

Relevant work on the regulation and the signals that induce the *S.* Typhi CRISPR-Cas system has been reported by our group (Hernández-Lucas et al., 2008; Medina-Aparicio et al., 2011, 2017). However, its biological function remains to be determined. In this regard, the low number of spacers in the CRISPR sequences, as well as their non-homology with bacteriophage and plasmid sequences, suggest that the CRISPR-Cas system does not provide an immune function in *Salmonella*. However, the conserved genetic organization of the *cas* genes in different *Salmonella* serovars is consistent with the system having a biological function in these bacteria (Touchon and Rocha, 2010; Shariat et al., 2015). In this work, it is reported that, in *S.* Typhi, CRISPR-Cas positively regulates OmpR, a two-component system regulator that induces the synthesis of the OmpC, OmpF, and OmpS2 porins. Additionally, it is demonstrated that the CRISPR-Cas system is involved in the resistance to bile salts and biofilm formation in *S.* Typhi.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Culture Conditions

The bacterial strains and plasmids used in this work are listed in **Supplementary Table S1**. *Salmonella* Typhi IMSS-1 (Puente et al., 1987) and *Escherichia coli* strains were grown aerobically at 37°C in LB (10 g tryptone, 5 g yeast extract, and 10 g NaCl per liter), MA (7 g nutrient broth, 1 g yeast extract, 2 ml glycerol, 3.75 g K_2HPO_4, and 1.3 g KH_2PO_4 per liter; Kawaji et al., 1979) or N-MM media [0.37 g KCl, 0.99 g $(NH_4)_2SO_4$, 0.087 g K_2SO_4, 0.14 g KH_2PO_4, 0.019 g $MgCl_2$, 1 g casamino acids, 5 ml glycerol, and 100 mM of Tris-HCl (pH 7.5) per liter] (Deiwick et al., 1999). When required, the following antibiotics were added: kanamycin (Km), 30 μg ml^{-1}; tetracycline (Tc), 12 μg ml^{-1}, and ampicillin (Ap), 200 μg ml^{-1}.

DNA Manipulations

Plasmid and genomic DNA isolations were carried out according to published protocols (Sambrook et al., 1989). Primers for PCR amplifications were provided by the Oligonucleotide Synthesis Facility at our Institute (**Supplementary Table S2**). Restriction enzymes, ligase, nucleotides, and polymerases were acquired from New England Biolabs, Invitrogen, or Thermo Scientific. For sequencing, double-stranded DNA was purified with the High Pure Plasmid Isolation Kit (Roche) and sequenced with an automatic Perkin Elmer/Applied Biosystems 377-18 system.

Site-Directed Mutagenesis

The *Salmonella* mutants were obtained by the one-step non-polar mutagenesis procedure (Datsenko and Wanner, 2000). The target

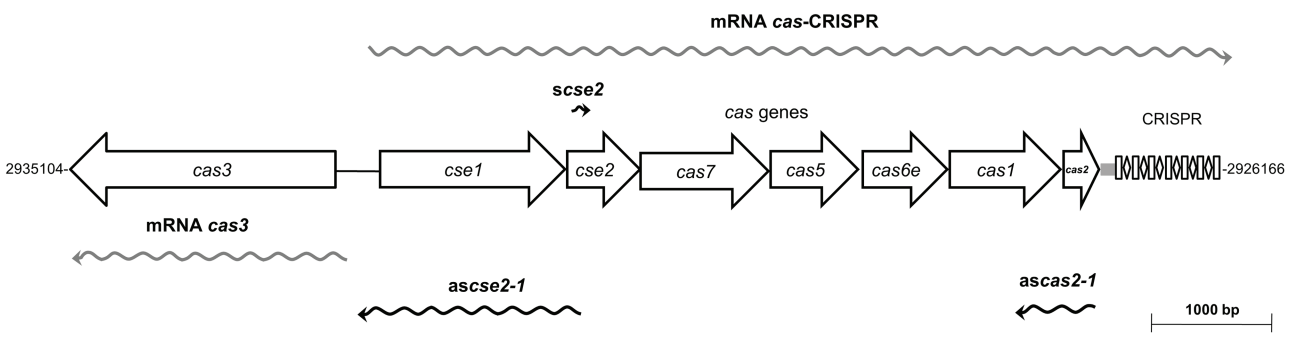


FIGURE 1 | Transcriptional organization of the Type I-E CRISPR-Cas system of *Salmonella enterica* serovar Typhi IMSS-1. The CRISPR-Cas system is composed of eight *cas* genes (*cas3* and *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*), a leader of 84 bp and the CRISPR1 array, containing seven 29-bp repeats and six 32-bp spacers. Five transcriptional units are transcribed from this locus, wavy gray arrows represent mRNAs of *cas*-CRISPR operon and *cas3*, whereas the RNA s*cse2*, the antisense as*cse2-1* and as*cas2-1* are shown as wavy black arrows. The *Salmonella* Typhi ΔCRISPR-*cas* strain is devoided of the entire Type I-E CRISPR-Cas system (from *cas3* to CRISPR locus).

gene was replaced with selectable antibiotic resistance gene markers. The resistance cassette was removed using the pCP20 plasmid. Each mutation was further characterized by sequencing to verify the authenticity of the deletion.

Construction of Transcriptional Reporter Fusions

For transcriptional *cat* constructs, oligonucleotides (see **Supplementary Table S2**) were designed to amplify DNA fragments of different lengths from the *ompC*, *ompF*, *ompS2*, and *ompR* regulatory regions. PCR products were double-digested with *Bam*HI-*Kpn*I and ligated into pKK232-8 or pKK232-9 (**Supplementary Table S1**), which contain the promoterless *cat* gene. All constructs were sequenced to verify the correct DNA sequence of the PCR fragments.

CAT Assays

To determine the expression of the *cat* reporter gene mediated by the *S.* Typhi promoters, chloramphenicol acetyltransferase (CAT) assays were performed according to a previously published protocol (Martínez-Laguna et al., 1999). Briefly, *S.* Typhi strains harboring the reporters were grown in N-MM or MA to different optical densities (OD), and the latter medium was supplemented when required with Ap and Km, with or without IPTG (isopropyl-β-D-thiogalactopyranoside; 50 μM). Cells were harvested, centrifuged, washed with 0.8 ml of TDTT buffer (50 mM Tris-HCl, 30 μM DL-dithiothreitol, and pH 7.8), resuspended in 0.5 ml of TDTT, and sonicated on ice for 10-s intervals with 10-s rest periods until the extract was clear. The homogenate was centrifuged at 12,000 *g*/15 min, and the supernatant used for activity measurement. For CAT assays, 5 μl of each extract were added in duplicate to a 96-well enzyme-linked immunosorbent assay (ELISA) plate, followed by the addition of 0.2 ml of a reaction mixture containing 1 mM DTNB [5,5'-dithiobis (2-nitrobenzoic acid)], 0.1 mM acetyl-coenzyme A (acetyl-CoA), and 0.1 mM chloramphenicol in 0.1 M Tris-HCl, pH 7.8. The absorbance at 412 nm was measured every 5 s for 5 min using a Ceres 900 scanning auto reader and microplate workstation. The protein concentration of the cell extracts was obtained using the bicinchoninic acid (BCA) protein assay reagent (Pierce). Protein values and the mean rate of product formation by CAT were used to determine CAT-specific activity as micromoles per minute per milligram of protein.

Preparation of Crude Cell Extracts for Two-Dimensional Gel Electrophoresis

Salmonella Typhi IMSS-1 and *S.* Typhi Δ*cas*-CRISPR harboring plasmid pFM*TrcleuO*-50 were grown in MA medium supplemented with Ap and IPTG (50 μM) to an optical density of 0.6 at 595 nm (OD_{595}). *Salmonella* cultures (100 ml) were pelleted and washed with 1X phosphate-buffered saline (PBS). Cellular proteins were obtained by sonication at 24 kHz for 1 min in the on position and 1 min in the off position, for five cycles at 4°C using a Vibra Cell (Sonics, United States), in the presence of a protease inhibitor (Complete tablets; Roche Diagnostics GmbH, Mannheim, Germany). To further limit proteolysis, protein isolation was performed using phenol extraction (Hurkman and Tanaka, 1986). To solubilize proteins and to obtain completely denatured and reduced proteins, pellets were dried and resuspended as previously reported (Encarnación et al., 2005). Prior to electrophoresis, samples were mixed with 7 M urea, 2 M thiourea, 4% 3-[(3-choloamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS; Roche Diagnostics GmbH, Germany), 2 mM tributylphosphine, 2% ampholytes, and 60 mM dithiothreitol.

Two-Dimensional Gel Electrophoresis

Methods used for sample preparation, analytical two-dimensional gel electrophoresis (2-DGE), image analysis, and preparative 2-DGE were described previously (Encarnación et al., 2003). pH gradients were determined using a two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis standard (Sigma, United States). For isoelectric focusing, 500 μg of total proteins were loaded. All gel experiments were repeated at least two times.

In-Gel Digestion and Mass Spectrometry-Based Identification of Proteins

Selected spots from Coomassie blue-stained preparative one- or two-dimensional gels were excised manually and frozen at −70°C until use. Samples were prepared for mass spectrum analysis using a slight modification of a previously described procedure (Encarnación et al., 2005). Protein spots were destained, reduced, alkylated, and digested with trypsin (Promega, Madison, WI). Before the mass spectra of the peptide mixtures were obtained, the mixtures were desalted using a C_{18} Zip Tip (Millipore, Bedford, MA) according to the manufacturer's recommendations. Mass spectra were determined using a Bruker Daltonics Autoflex (Bruker Daltonics, Billerica, MA) operated in the delayed extraction and reflectron mode. Spectra were externally calibrated using a peptide calibration standard (Bruker Daltonics 206095). Peptide mixtures were analyzed using a saturated solution of alpha-cyano-4-hydroxycinnamic acid in 50% acetonitrile-0.1% trifluoroacetic acid. Peak lists of the tryptic peptide masses were generated and searched against the NCBInr databases using the Mascot search program (Matrix Science, London, United Kingdom).[1]

Preparation of Outer Membrane Proteins

Outer Membrane Proteins (OMPs) were isolated from *S.* Typhi IMSS-1 strains grown in N-MM to an OD_{595} of 0.6 and 1.3 according to previous protocols (Puente et al., 1995). Fifteen milliliter of each culture was harvested and centrifuged at 5,000 *g* for 10 min at 4°C. Cells were resuspended in 500 μl of 10 mM Na_2HPO_4 buffer (pH 7.2) and sonicated on ice until the suspensions were clear. Intact cells and debris were eliminated by centrifugation (15,000 *g*) for 2 min, and the supernatants were transferred to clean microcentrifuge tubes and membrane fractions were pelleted by centrifugation at

[1]http://www.matrixscience.com

12,000 *g* for 1 h at 4°C. Inner membrane proteins were solubilized by resuspension in 500 µl of 10 mM Na_2HPO_4 buffer, pH 7.2, containing 2% Triton X-100 for 30 min at 37°C. After incubation, the samples were centrifuged at 12,000 *g* for 1 h at 4°C. The remaining outer membrane insoluble fraction was washed with 500 µl of 10 mM Na_2HPO_4, pH 7.2, centrifuged at 12,000 *g* for 1 h at 4°C, and finally resuspended in 50 µl 1X PBS, pH 7.4. OMP concentrations were determined by BCA assay (Thermo), and 15 µg of each sample was analyzed by SDS-12% polyacrylamide gel electrophoresis. One-dimensional OMP gels were visualized by staining with Coomassie brilliant blue.

Western Blotting

For western blot experiments, *S.* Typhi wild-type strain and its derivatives were grown in N-MM to OD_{595} of 1.0 or MA medium to an OD_{595} of 0.6. The cultures were supplemented, when required, with Ap and IPTG (50 µM). Fifteen milliliter of each culture was harvested and centrifuged at 5,000 *g* for 8 min. The pellets were resuspended in 600 µl of 1X PBS and sonicated on ice for 12 min at intervals of 10-s with 5-s rest. Total protein concentration was determined by BCA assay (Thermo), and 80 µg of each sample was loaded on a 10% SDS polyacrylamide gel. Following electrophoresis, proteins were transferred to 0.45-µm-pore-size polyvinylidene difluoride membranes (Immobilon; Millipore) using the Trans-Blot SD system (Bio-Rad) according to a previously described procedure (Guadarrama et al., 2014). Membranes were blocked with 10% non-fat milk and incubated with anti-OmpR or anti-GroEL (StressGen) polyclonal antibodies. Then, they were washed with 1X PBS, 0.1% Tween 20. Immunodetection was performed with a 1:10,000 dilution of horseradish peroxidase-conjugated Anti-Rabbit antibody (Pierce) for polyclonal antibodies, and the Western Lightning Plus-ECL Chemiluminescence Reagent Kit (PerkinElmer). The membranes containing the proteins were exposed to Carestream X-OMAT LS films.

Growth Evaluation in 5% Sodium Deoxycholate

Salmonella Typhi wild-type and the different mutant strains were grown 24 h in LB plates at 37°C. A bacterial colony was inoculated in liquid LB broth (5 ml) and grown for 16 h at 37°C/200 rpm. Then, 50 ml of LB broth supplemented with 5% sodium deoxycholate (Sigma Chemical, St. Louis, MO) were inoculated with the pre-inoculum to give an initial OD at 595 nm of 0.02. The cultures were incubated at 37°C/200 rpm during 15 h with OD_{595} measurements being done every 2 h.

Microtiter Dish Biofilm Formation Assay

The quantification of biofilm formation was performed following a previous established protocol (O'Toole, 2011). Briefly, bacterial cells were grown overnight in LB broth (5 ml) at 37°C/200 rpm. Cells were diluted 1:100 in fresh LB without NaCl for stimulates biofilm production. One hundred microliter of this dilution was added per well in a 96-well polystyrene microtitre plate (Costar Cat. No. 3599, flat bottom with lid). Six replicate wells were prepared for each strain. Microtitre plates were incubated at 30°C for 24 h. Total bacterial growth was measured at OD_{600}, using a GloMax®-Muti Detection System (Promega). The planktonic cells were then discarded, and the plate was washed three times with water. The remaining biofilm was fixed with 200 µl per well of methanol (100%) and stained with a 0.2% solution of crystal violet in water. After incubation at room temperature for 10 min, the plates were rinsed three times with water. The dye was solubilized by adding 125 µl of 33% acetic acid to each well and incubated the microtiter plate at room temperature for 15 min. Finally, the OD_{560} was determined with the microplate reader. The amount of formed biofilm is reported as the ratio of the OD_{560}/OD_{600} values (Oropeza et al., 2015).

RESULTS

CRISPR-Cas Is Fundamental for the Synthesis of Major and Quiescent Outer Membrane Proteins in *Salmonella* Typhi

Studies on the regulation and the signals that induce the CRISPR-*cas* locus in *S.* Typhi are available (Hernández-Lucas et al., 2008; Medina-Aparicio et al., 2011, 2017). However, its biological function in this human pathogen remains to be determined. In this regard, previous results in *Francisella novicida* demonstrated that the CRISPR-Cas system is involved in the synthesis of outer membrane proteins (Sampson et al., 2013, 2014). Therefore, we obtained a strain devoid of *cas3*, the *cas3-cse1* intergenic region, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2*, and the CRISPR locus (the entire Type I-E CRISPR-Cas system, **Figure 1**), which was named as ΔCRISPR-*cas* (**Supplementary Table S1**). By electrophoretic profiles, the presence of the major outer membrane proteins OmpC, OmpF, and OmpA was detected in the wild-type strain; whereas in the isogenic *S.* Typhi strain devoid of CRISPR-*cas* locus, OmpC, and OmpF were not visualized (**Figure 2A**). To confirm these results, the transcriptional expression of *ompC* and *ompF* promoter regions was evaluated. Thus, the reporter plasmids pKK9/*ompC*-772 + 27 and pKK8/*ompF*-782 + 184 (**Supplementary Table S1**) were transformed into *S.* Typhi IMSS-1 wild type and, in the isogenic ΔCRISPR-*cas* strain, to perform CAT assays. The experiments showed that the transcriptional activity of the *ompC* and *ompF* regulatory regions were of 4,328 and 5,512 CAT units, respectively, in the wild-type strain. However, in the ΔCRISPR-*cas* strain the *ompC* and *ompF* activity decreased by 99 and 73%, respectively (**Figure 2B**). These data demonstrated that the CRISPR-Cas system is relevant for the expression of the major OmpC and OmpF porins in *S.* Typhi.

To determine the specific CRISPR-Cas genetic element involved in OmpC and OmpF regulation, a collection of individual *cas* mutants was generated, and porin profiles of these strains showed that Δ*cse1*, Δ*cas7*, and wild-type *S.* Typhi present a similar outer membrane protein profile. Nevertheless, in the absence of *cas5* and *cas2*, the amount of OmpC decreased substantially; whereas in the individual

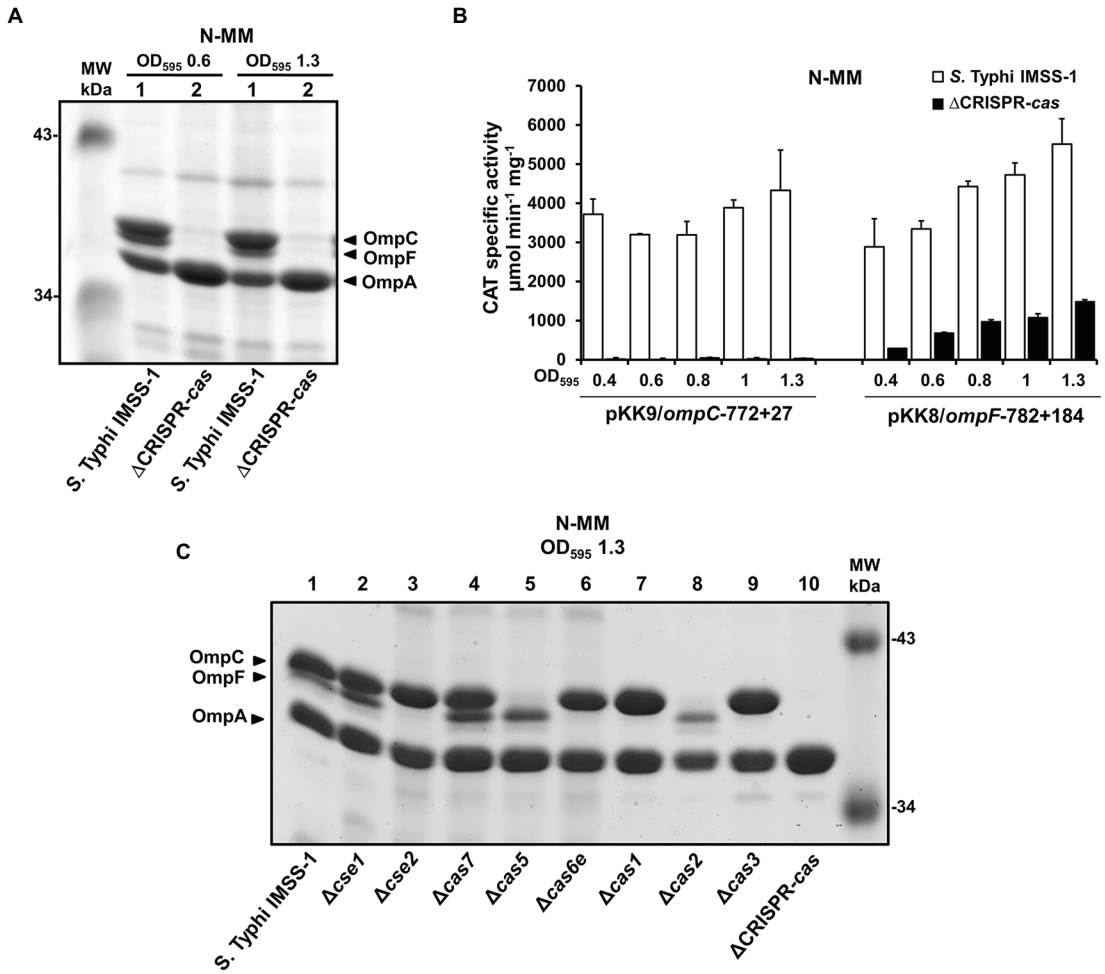
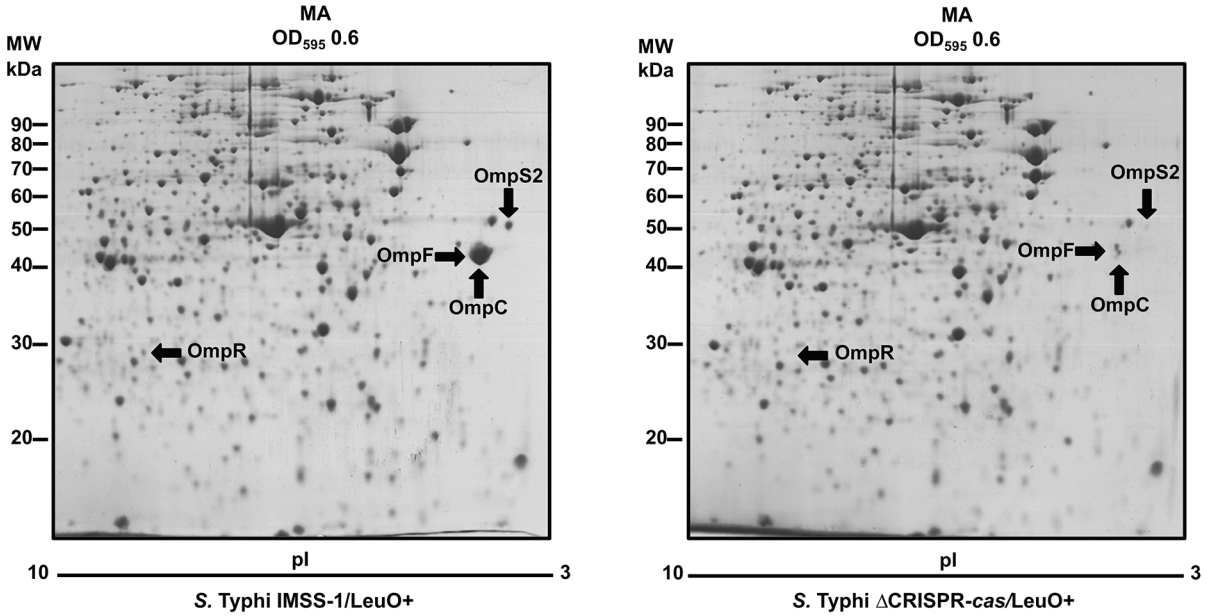


FIGURE 2 | CRISPR-Cas is fundamental for the synthesis of the major outer membrane proteins OmpF and OmpC in *Salmonella* Typhi. **(A)** Electrophoretic pattern of Coomassie brilliant blue-stained outer membrane protein preparations, separated by 0.1% SDS-15% PAGE. The bacterial strains *Salmonella* Typhi IMSS-1 wild type (lane 1) and *Salmonella* Typhi ΔCRISPR-*cas* (ΔCRISPR-*cas*, lane 2) were grown in N-MM to an OD_{595} of 0.6 and 1.3. The major OMPs, OmpC, OmpF, and OmpA are indicated with a black triangle. **(B)** Transcriptional profiles of *Salmonella* Typhi IMSS-1 and *Salmonella* Typhi ΔCRISPR-*cas* harboring plasmid pKK9/*ompC*-772 + 27 or pKK8/*ompF*-782 + 184 in N-MM. CAT-specific activities were measured at an OD_{595} of 0.4, 0.6, 0.8, 1.0, and 1.3. The values are the means ± standard deviations for three independent experiments performed in duplicate. **(C)** Electrophoretic pattern of Coomassie brilliant blue-stained outer membrane protein preparations, separated by 0.1% SDS-15% PAGE from *Salmonella* Typhi IMSS-1 wild type (lane 1), Δ*cse1* (lane 2), Δ*cse2* (lane 3), Δ*cas7* (lane 4), Δ*cas5* (lane 5), Δ*cas6e* (lane 6), Δ*cas1* (lane 7), Δ*cas2* (lane 8), Δ*cas3* (lane 9), and *Salmonella* Typhi ΔCRISPR-*cas* (ΔCRISPR-*cas*, lane 10) strains, grown in N-MM at OD_{595} of 1.3. The OmpC, OmpF, and OmpA porins are indicated with a black triangle. Molecular weight markers (MW) are indicated.

cse2, *cas6e*, *cas1* and *cas3* mutants the OmpF porin was not observed (**Figure 2C**). These data support the fundamental role of specific Cas proteins in the regulation of OmpC and OmpF major outer membrane proteins and also are in agreement with the results obtained from the deletion of the entire CRISPR-Cas locus, since this strain lacks *cas5*, *cas2*, *cse2*, *cas6e*, *cas1*, and *cas3*, which resulted in the absence of the two main porins in *S.* Typhi (**Figure 2A**).

To continue with the identification of more CRISPR-Cas dependent outer membrane proteins, and since the overexpression of LeuO induces quiescent porins, such as OmpS2 (Fernández-Mora et al., 2004), the induction of this protein was evaluated in the absence of CRISPR-Cas. *Salmonella* Typhi IMSS-1 harboring plasmid pFM*TrcleuO*-50 and *S.* Typhi ΔCRISPR-*cas* containing pFM*TrcleuO*-50 were grown to an OD_{595} of 0.6 in MA medium supplemented with IPTG (50 μM), and 2-DGE profiles were obtained with these cultures. The results showed the presence of OmpS2 in the wild-type strain. However, in the absence of the CRISPR-*cas* locus, OmpS2 decreased its expression by 99% (**Figure 3A**). Even more, the expression of a transcriptional fusion of the 5' intergenic region of *ompS2* (pKK9/*ompS2*-482 + 77, **Supplementary Table S1**), upon induction of the LeuO regulator at various points in the growth curve, was essentially abolished in the ΔCRISPR-*cas* as compared with the wild-type strain (**Figure 3B**). Therefore, CRISPR-Cas is also fundamental for OmpS2 expression mediated by LeuO.

A

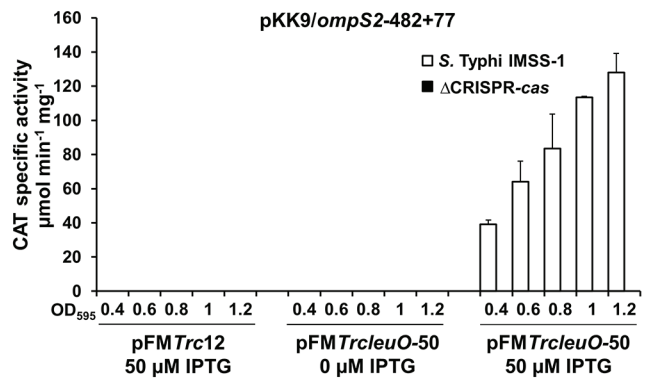


CRISPR-Cas-regulated proteins in *Salmonella enterica* serovar Typhi identified by MALDI-TOF

Spot	Protein	Abundance (*S.* Typhi IMSS-1/ *S.* Typhi ΔCRISPR-*cas*)[a]	LeuO induction	Matching database protein		pI [b]	Mol wt[b]
				Organism	Accession no.		
OmpS2	Porin	135.4/0.7	+	*S.* Typhi	STY1649	4.3	51
OmpC	Porin	5475/15.9	-	*S.* Typhi	STY2493	4.63	41.25
OmpF	Porin	5475/27	-	*S.* Typhi	STY1002	4.68	40.11
OmpR	Transcriptional regulator	21.2/0	-	*S.* Typhi	STY4294	6.9	27.3

[a] Protein spot intensity (OD) of *S.* Typhi containing pFM*TrcleuO*-50 (*S.* Typhi IMSS-1)/*S.* Typhi ΔCRISPR-*cas* containing pFM*TrcleuO*-50 (*S.* Typhi ΔCRISPR-*cas*).
[b] The values were obtained in three independent experiments.

B



C

***ompS2* transcriptional expression**

S. Typhi strain	CAT specific activity (μmol min^{-1} mg^{-1})	
	pFM*TrcleuO*-50 0 μM IPTG	pFM*TrcleuO*-50 50 μM IPTG
IMSS-1 + pKK9/*ompS2*-482+77	<dl	71.7 ± 4.8
Δ*cse1* + pKK9/*ompS2*-482+77	<dl	43.1 ± 16.1
Δ*cse2* + pKK9/*ompS2*-482+77	<dl	<dl
Δ*cas7* + pKK9/*ompS2*-482+77	<dl	76.6 ± 19.0
Δ*cas5* + pKK9/*ompS2*-482+77	<dl	<dl
Δ*cas6e* + pKK9/*ompS2*-482+77	<dl	<dl
Δ*cas1* + pKK9/*ompS2*-482+77	<dl	<dl
Δ*cas2* + pKK9/*ompS2*-482+77	<dl	<dl
Δ*cas3* + pKK9/*ompS2*-482+77	<dl	<dl

FIGURE 3 | CRISPR-Cas is essential for the synthesis of Omps2, OmpC, and OmpF in *Salmonella* Typhi. **(A)** Proteomic profiles of *Salmonella* Typhi total protein extracts. *Salmonella* Typhi IMSS-1 wild type and *Salmonella* Typhi ΔCRISPR-*cas* containing pFM*TrcleuO*-50 were grown in MA to an OD_{595} of 0.6. Cultures were

(Continued)

FIGURE 3 | supplemented with 50 μM IPTG. The labeled spots were excised and identified using MALDI-TOF. Below the 2-DGE the CRISPR-Cas regulated proteins in *Salmonella* Typhi identified by MALDI-TOF are shown. **(B)** *ompS2* transcriptional activities. *Salmonella* Typhi IMSS-1 (white bars) and *Salmonella* Typhi ΔCRISPR-*cas* [ΔCRISPR-*cas*, black bars (values below the detection limit)] harboring plasmid pFM*TrcleuO*-50 or pFM*Trc*12 were independently transformed with pKK9/*ompS2*-482 + 77. The strains were grown in MA medium and CAT-specific activity was measured at OD_{595} of 0.4, 0.6, 0.8, 1.0, and 1.2. **(C)** Expression profiles of *Salmonella* Typhi IMSS-1, Δ*cse1*, Δ*cse2*, Δ*cas7*, Δ*cas5*, Δ*cas6e*, Δ*cas1*, Δ*cas2*, and Δ*cas3* strains containing pFM*TrcleuO*-50 and pKK9/*ompS2*-482 + 77 plasmids. The strains were grown in MA medium at OD_{595} of 1.0. The values are the means ± standard deviations for three independent experiments performed in duplicate; <dl (<detection limit) represents values between 0 and 10 CAT units.

To determine the specific *cas* genetic element involved in OmpS2 regulation, the individual *cas* mutants were transformed with the transcriptional CAT fusion containing the 5' regulatory region of *ompS2* and plasmid pFM*TrcleuO*-50 for overexpressing LeuO. The expression results showed that *ompS2* activity mediated by LeuO depends on *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3*, since in the absence of each of these genetic elements *ompS2* was not transcribed (**Figure 3C**). These results indicated that the majority of Cas proteins, with exception of Cse1 and Cas7, are essential for the synthesis of the quiescent porin OmpS2. Thus, the presence of CRISPR-Cas cluster is essential for the synthesis of major and quiescent porins in *S.* Typhi.

The *Salmonella* Typhi CRISPR-Cas System Is Involved in the Expression of the Porin Master Regulator OmpR

The results mentioned above showed that CRISPR-Cas is involved in the synthesis of outer membrane proteins in *S.* Typhi. Interestingly, in the 2-DGE image shown in **Figure 3A**, a small spot of 27.3 kDa was absent in the ΔCRISPR-*cas*, and the mass spectrometry (MS) results of the same spot from *S.* Typhi IMSS-1, demonstrated that it corresponded to OmpR. To define whether CRISPR-Cas is involved in the control of the gene for this two-component system regulator, we evaluated its transcriptional expression in the *S.* Typhi wild type and in a ΔCRISPR-*cas*. The results showed that *ompR* displayed 941 CAT units in the wild-type strain, and the activity decreased by 60% in the CRISPR-Cas deficient *S.* Typhi strain (**Figure 4A**). Previously, it was demonstrated that *ompR* contains two promoters (Villarreal et al., 2014). To define whether the *ompR*P1 or *ompR*P2 promoters are under CRISPR-Cas control, the transcriptional activity of each promoter in the wild-type strain and in the ΔCRISPR-*cas* was evaluated. The transcriptional results showed *ompR*P2 activity values of 248 and 279 CAT units in the wild type and in the ΔCRISPR-*cas* isogenic strain, respectively (**Figure 4A**). Thus, CRISPR-Cas is not involved in *ompR*P2 promoter control. However, the activities obtained with *ompR*P1 were 120 and 37 CAT units in the wild type and in the ΔCRISPR-*cas*, respectively (**Figure 4A**). Therefore, CRISPR-Cas is involved in the regulation of the *ompR*P1 promoter to induce *ompR* expression.

To validate the results obtained and to determine the Cas proteins involved in *ompR*P1 genetic control, individual *cas* mutants were transformed with CAT fusions containing either the *ompR*P2 (pKK8/*ompR*P2-383-133) or the *ompR*P1 (pKK8/*ompR*P1-134-1) promoters. The activity results showed that in the *S.* Typhi wild type as well as in individual *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* null mutants, the *ompR*P2 promoter expression was similar, supporting the notion that the Cas proteins are not implicated in its regulation (**Figure 4B**). In the case of the *ompR*P1 promoter, its genetic activity in the individual *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* deficient strains was considerably reduced, compared with the CAT units obtained in the wild-type strain and in the *cse1* and *cas7* mutants (**Figure 4C**). The data support the proposal that the Cas proteins involved in *ompR*P1 promoter regulation correspond to Cse2, Cas5, Cas6e, Cas1, Cas2, and Cas3; whereas Cse1 and Cas7 are not implicated in *ompR*P1 induction.

To determine whether the reduction of *ompR*P1 promoter activity in the ΔCRISPR-*cas*, as well as in each *cas* individual mutant, has an effect on the synthesis of OmpR, western blot experiments were performed. The wild-type *S.* Typhi, the ΔCRISPR-*cas*, as well as the individual *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, and *cas2* deletion mutants were grown in N-MM to an OD_{595} of 1.0. Total crude cell protein extracts were transferred to membranes and probed using anti-OmpR polyclonal antibody. The western blot results showed a prominent OmpR band of 27.3 KDa in the wild-type strain and in the *cse1* and *cas7* individual mutants; whereas in the ΔCRISPR-*cas* strain, and in the individual *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* mutants the OmpR protein was absent (**Figure 4D**).

These results explain the lack of OmpC and OmpF in the corresponding *cas* deficient strains (**Figure 2C**), since it is well-known that OmpR binds to their regulatory regions to promote their expression (Yoshida et al., 2006). Therefore, *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* genes are fundamental for OmpR expression, whereas *cse1* and *cas7* are not involved in OmpR regulation, demonstrating that specific *cas* genes are necessary for OmpR production to control porin synthesis. Moreover, complementation of the *S.* Typhi ΔCRISPR-*cas* with the *ompR* gene on a plasmid restored the presence of OmpC and OmpF porins in this strain (**Figure 4E**), further supporting the notion that the deletion of the entire CRISPR-*cas* loci results in the lowering of the expression of the OmpR regulator and thus porin expression.

With respect to the OmpS2 quiescent porin, it is well accepted that LeuO counteracts the negative effect of H-NS on the *ompS2* promoter, upon which OmpR binds to its regulatory region promoting *ompS2* expression (Fernández-Mora et al., 2004). Thus, it was determined whether the OmpR protein was produced in the *S.* Typhi strains that overexpress LeuO. Western blot experiments demonstrated the presence of OmpR in *S.* Typhi IMSS-1 wild-type strain harboring the pFM*TrcleuO*-50 plasmid. However, OmpR was not detected in the *S.* Typhi Δ*cas*-CRISPR mutant overexpressing LeuO (**Figure 4F**). Thus, OmpS2 was not visualized in the 2-DGE of this strain (**Figure 3A**) because of the lack of the two-component system regulator OmpR.

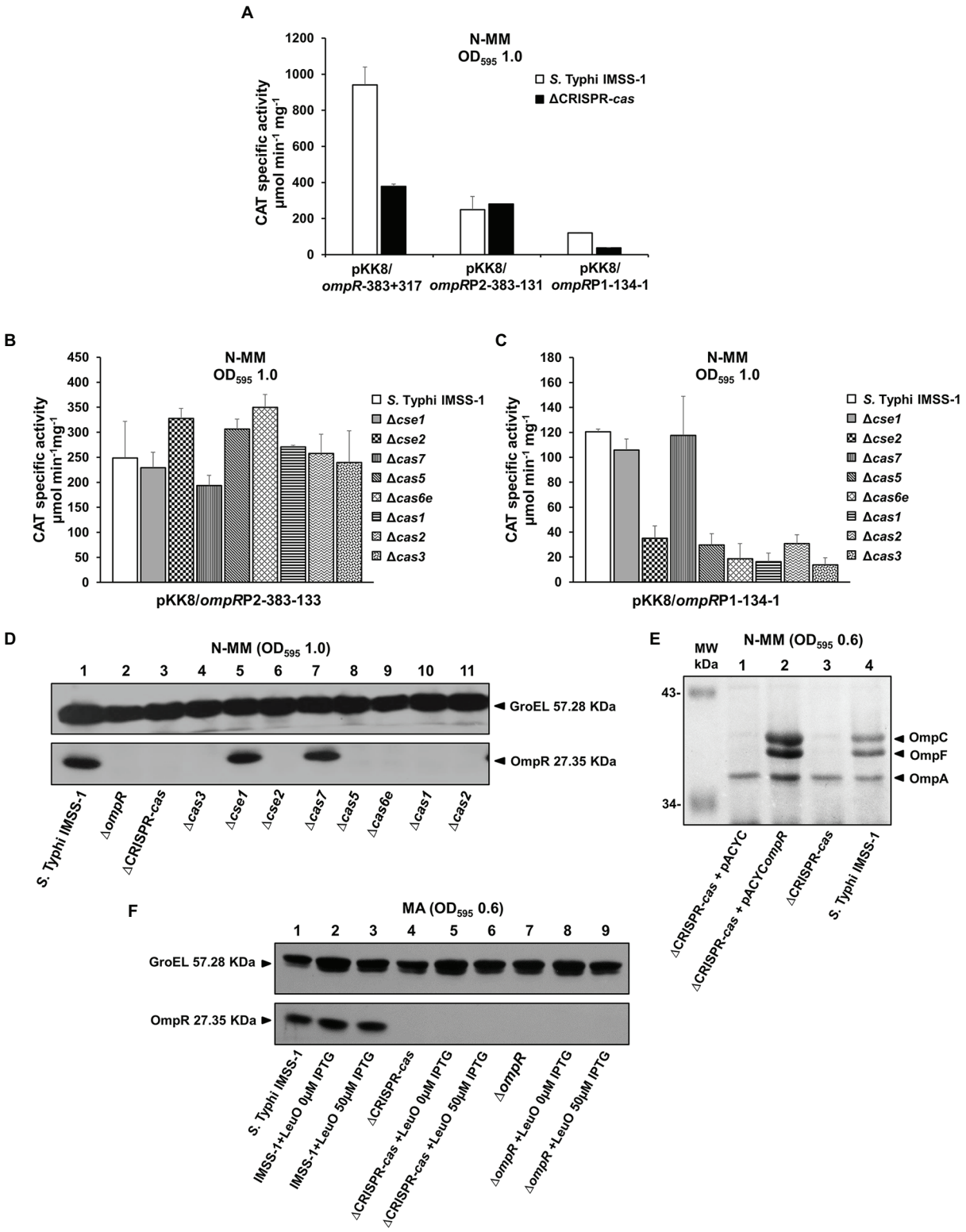


FIGURE 4 | CRISPR-Cas is involved in the genetic control of OmpR. **(A)** Transcriptional profiles of *Salmonella* Typhi IMSS-1 and *Salmonella* Typhi IMSS-1 ΔCRISPR-*cas* (ΔCRISPR-*cas*) harboring plasmids pKK8/*ompR*-383 + 317, pKK8/*ompR*P2-383-133 (*ompR*P2) or pKK8/*ompR*P1-134-1 (*ompR*P1), grown in

(Continued)

FIGURE 4 | N-MM. CAT-specific activity was measured at an OD_{595} of 1.0. The values are the means ± standard deviations of three independent experiments performed in duplicate. The transcriptional expression of pKK8/*ompR*P2-383-133 **(B)** and pKK8/*ompR*P1-134-1 **(C)** was also evaluated in *Salmonella* Typhi IMSS-1, Δ*cse1*, Δ*cse2*, Δ*cas7*, Δ*cas5*, Δ*cas6e*, Δ*cas1*, Δ*cas2*, and Δ*cas3* strains grown in N-MM. The samples were collected at OD_{595} of 1.0. The values are the means ± standard deviations for three independent experiments performed in duplicate. **(D)** Western blot using anti-OmpR polyclonal antibody and total proteins from *Salmonella* Typhi IMSS-1 wild type (lane 1), *Salmonella* Typhi Δ*ompR* (Δ*ompR*, lane 2), *Salmonella* Typhi ΔCRISPR-*cas* (ΔCRISPR-*cas*, lane 3), Δ*cas3* (lane 4), Δ*cse1* (lane 5), Δ*cse2* (lane 6), Δ*cas7* (lane 7), Δ*cas5* (lane 8), Δ*cas6e* (lane 9), Δ*cas1* (lane 10), and Δ*cas2* (lane 11) strains were grown in N-MM at OD_{595} of 1.0. **(E)** Electrophoretic pattern of Coomassie brilliant blue-stained outer membrane protein preparations, separated by 0.1% SDS-15% PAGE of *Salmonella* Typhi ΔCRISPR-*cas* + pACYC (lane 1), *Salmonella* Typhi ΔCRISPR-*cas* + pACYC*ompR* (lane 2), *Salmonella* Typhi ΔCRISPR-*cas* (lane 3), and *Salmonella* Typhi IMSS-1 wild type (lane 4), grown in N-MM to an OD_{595} of 0.6. The major OMPs: OmpC, OmpF, and OmpA are indicated with a black triangle. **(F)** Western blot performed with anti-OmpR polyclonal antibody and total proteins from *Salmonella* Typhi IMSS-1 (lane 1), *Salmonella* Typhi IMSS-1 + pFM*TrcleuO*-50 (0 μM IPTG; lane 2), *Salmonella* Typhi IMSS-1 + pFM*TrcleuO*-50 (50 μM IPTG; lane 3), *Salmonella* Typhi IMSS-1 ΔCRISPR-*cas* (lane 4), *Salmonella* Typhi IMSS-1 ΔCRISPR-*cas* + pFM*TrcleuO*-50 (0 μM IPTG; lane 5), *Salmonella* Typhi IMSS-1 ΔCRISPR-*cas* + pFM*TrcleuO*-50 (50 μM IPTG; lane 6), *Salmonella* Typhi IMSS-1 Δ*ompR* (lane 7), *Salmonella* Typhi IMSS-1 Δ*ompR* + pFM*TrcleuO*-50 (0 μM IPTG; lane 8), *Salmonella* Typhi IMSS-1 Δ*ompR* + pFM*TrcleuO*-50 (50 μM IPTG, lane 9). All the *Salmonella* Typhi bacterial strains were grown in MA to an OD_{595} of 0.6. GroEL was used as protein loading control. The proteins visualized are indicated with black triangles.

In conclusion, the results obtained here showed that the CRISPR-Cas system acts hierarchically on the *ompR*P1 promoter to induce OmpC, OmpF, or OmpS2 synthesis in *S.* Typhi.

The *Salmonella* Typhi *cas* Genes Are Involved in Sodium Deoxycholate Resistance and Biofilm Formation

In this report, we have shown that the *S.* Typhi *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* genes are involved in porin synthesis through the regulation of the *ompR* gene which codes for the OmpR transcriptional regulator. In previous studies, it has been demonstrated that *ompR* is involved in virulence, sodium deoxycholate resistance, biofilm formation, the production of flagella, and curli (Pickard et al., 1994; Shin and Park, 1995; Vidal et al., 1998; Cameron and Dorman, 2012; Villarreal et al., 2014). Therefore, we evaluated whether the *cas* genes are involved in some of these biological processes. Growth rate experiments of *S.* Typhi IMSS-1 and the *cas* individual deleted strains were performed in LB broth supplemented with 5% of the human bile salt sodium deoxycholate. The results showed that the wild-type *S.* Typhi strain grew in this condition, reaching an OD_{595} of 0.86 after 15 h. However, growth of the *Δcas5*, Δ*cas2*, and *Δcas*-CRISPR mutant strains was impaired in the presence of this bile salt (**Figure 5A**) since their OD_{595} were of 0.43, 0.54, and 0.18, respectively, after 15 h of incubation. Remarkably, these strains did not express the OmpC porin, which was previously shown to be determinant for allowing *S.* Typhi to proliferate in the presence of sodium deoxycholate (Villarreal et al., 2014). The growth rate of Δ*cas6e*, Δ*cas1*, and Δ*cas3* was similar to that observed with the wild-type strain (**Figure 5A**), consistent with the presence of the OmpC porin in these mutants. The same experiment was performed with Δ*ompR*, Δ*ompC*, Δ*ompF*, and Δ*ompS2* strains. As expected, the *ompR* and *ompC* mutants were also impaired in their growth in 5% sodium deoxycholate (OD_{595} = 0.5 and 0.33, respectively; **Figure 5B**), as previously reported (Villarreal et al., 2014); whereas the Δ*ompF* and Δ*ompS2* mutant strains grew like the *S.* Typhi IMSS-1 wild type.

Additionally, we also evaluated the biofilm formation ability of the *S.* Typhi IMSS-1 wild type, and of the Δ*cse2*, Δ*cas5*, Δ*cas6e*, Δ*cas1*, Δ*cas2*, Δ*cas3*, *Δcas*-CRISPR, Δ*ompR*, Δ*ompC*, Δ*ompF*, and Δ*ompS2* mutant strains. The experiments showed that the wild type produced moderate biofilm (0.26 OD_{560}/OD_{600} ratio). However, the *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, *cas3*, and *cas*-CRISPR mutants displayed an increased biofilm formation (OD_{560}/OD_{600} ratio of 1.06, 1.04, 0.95, 1.04, 1.17, 0.90, and 0.50, respectively; **Figure 6**). These results suggest that the CRISPR-Cas system negatively regulates genes involved in biofilm production, i.e., that the absence of *cas* genes allows the expression of factors that increase the ability of *S.* Typhi to form biofilm.

The *ΔompR* mutant showed an increased biofilm formation (0.77), as compared to the values obtained with the wild type (0.26; **Figure 6**). However, strains carrying deletions either in the *ompC*, *ompF*, or the *ompS2* genes presented a slightly decreased biofilm formation (OD_{560}/OD_{600} ratio of 0.15, 0.14, and 0.18, respectively; **Figure 6**). Therefore, the biofilm production was independent of the individual absence of the OmpC, OmpF, or OmpS2 porins. Interestingly, it has been demonstrated that *ompR* mutants in *Salmonella enteritidis*, *Salmonella pullorum*, *E. coli*, and *Yersinia enterocolitica* presented a decreased biofilm formation ability (Dong et al., 2011; Lu et al., 2012; Samanta et al., 2013; Meng et al., 2019), suggesting that the pathway toward regulating biofilm synthesis is different in *S.* Typhi.

The data shown are consistent with the notion that the CRISPR-Cas system is relevant for *S.* Typhi virulence, since this pathogen needs to survive the presence of bile salts in the gut and gallbladder, as well as to persist inside the gallbladder, where the biofilm formation is relevant (Crawford et al., 2010; Gonzalez-Escobedo et al., 2011; Spector and Kenyon, 2012).

DISCUSSION

The results presented here showed that the Cse2, Cas5, Cas6e, Cas1, Cas2, and Cas3 proteins *via* the positive regulation of the two-component regulator OmpR, have a role in the major and quiescent outer membrane protein synthesis, since they control OmpC, OmpF, and OmpS2. Due to the fact that only a few transcriptional factors have been implicated in the control of *ompR* in *Salmonella*, such as LtrR, H-NS, and OmpR (autoregulation; Bang et al., 2002; Villarreal et al., 2014), the data obtained contribute to the understanding of the regulatory network that controls the activity of this master regulator.

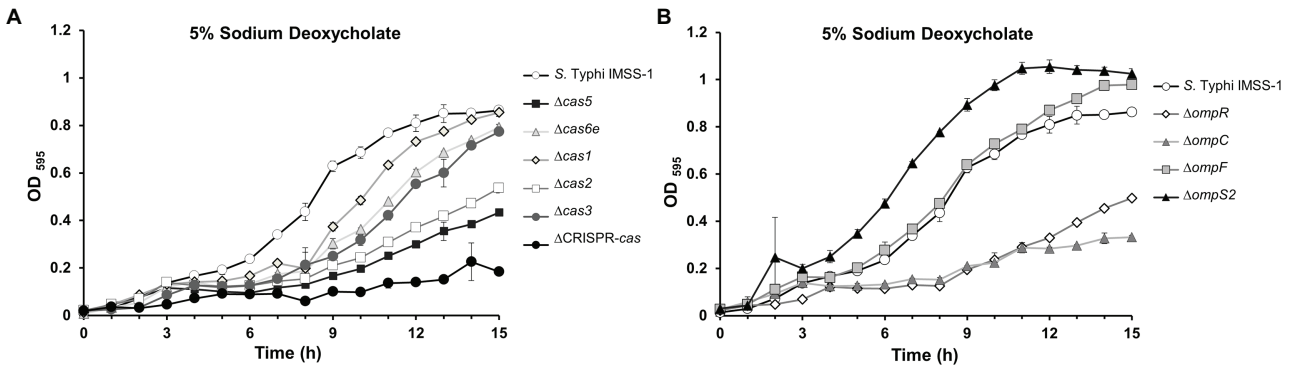


FIGURE 5 | The CRISPR-Cas system is required for resistance to the bile salt sodium deoxycholate in *Salmonella* Typhi. **(A)** Growth kinetics in sodium deoxycholate of *Salmonella* Typhi IMSS-1 wild type (empty circle), Δ*cas5* (black square), Δ*cas6e* (gray triangle), Δ*cas1* (gray diamond), Δ*cas2* (empty square), Δ*cas3* (gray circle), and ΔCRISPR-*cas* (black circle). **(B)** Growth kinetics in sodium deoxycholate of *Salmonella* Typhi IMSS-1 wild type (empty circle), Δ*ompR* (empty diamond), Δ*ompC* (gray triangle), Δ*ompF* (gray square), and Δ*ompS2* (black triangle). For both kinetics, the bacterial strains were grown in LB broth supplemented with 5% sodium deoxycholate at 37°C. The growth was monitored by OD_{595}. Three independent experiments were performed in duplicate and representative data are shown.

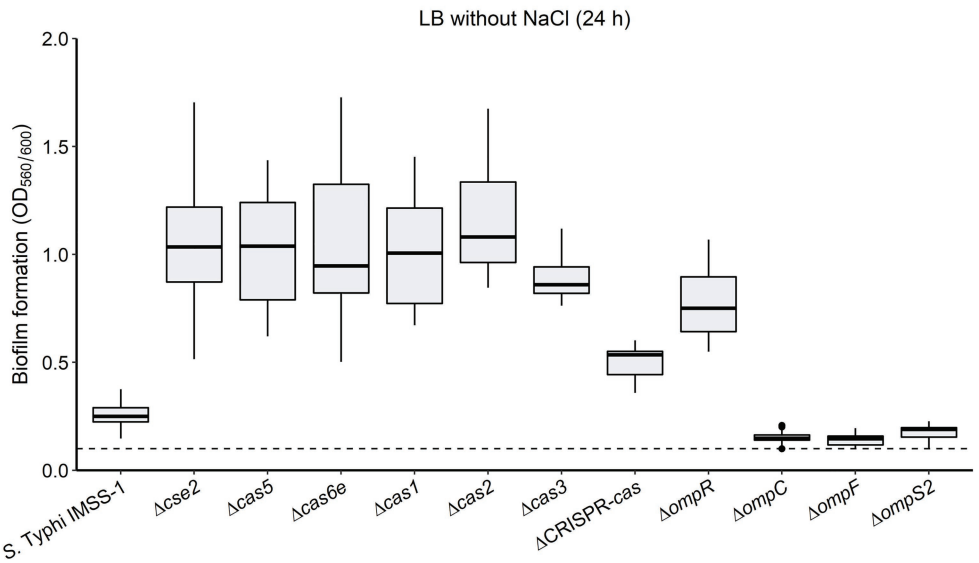


FIGURE 6 | The absence of the CRISPR-Cas system increases biofilm formation in *Salmonella* Typhi. The *Salmonella* Typhi IMSS-1 wild type, and Δ*cse2*, Δ*cas5*, Δ*cas6e*, Δ*cas1*, Δ*cas2*, Δ*cas3*, ΔCRISPR-*cas*, Δ*ompR*, Δ*ompC*, Δ*ompF*, and Δ*ompS2* mutant strains were grown in LB broth without NaCl and incubated at 30°C for 24 h to evaluate biofilm formation by crystal violet staining in microtitre dishes. The dotted line indicates a 0.1 background value and a bacterial strain produces biofilm when the values obtained are above 0.1. The box plot diagram represents the data obtained from three independent experiments.

The results also support the complex genetic regulation of porins (De la Cruz and Calva, 2010), since in the absence of *cas5* and *cas2*, OmpR becomes undetectable (**Figure 4D**), as does OmpC (**Figure 2C**), demonstrating the specific role of these *cas* genes on *ompR* regulation to mediate OmpC synthesis. Interestingly, the presence of OmpF was evident in these *cas* mutants, supporting the notion that OmpF is not only OmpR-dependent, and that other transcriptional factors are able to induce OmpF expression. In this sense, regulators, such as Lrp and CadC, are also involved in its positive control (Ferrario et al., 1995; Lee et al., 2007; De la Cruz and Calva, 2010). In contrast, in the individual *cse2*, *cas6e*, *cas1*, and *cas3* mutants the OmpF porin was not visualized (**Figure 2C**), and OmpR was not detected by western blot (**Figure 4D**), supporting the role of these genes in the control of *ompR* to promote OmpF synthesis. In these *cas* mutants, the presence of OmpC was observed, supporting the proposal that other regulators are able to induce OmpC synthesis. In this respect, the CpxRA and CadC transcriptional factors have been reported to positively regulate *ompC* (Batchelor et al., 2005; Lee et al., 2007; De la Cruz and Calva, 2010).

In *E. coli*, it is well-known that *ompR-envZ* comprises an operon, and a bioinformatic analysis using the Operon-mapper tool suggested that, in *S.* Typhi, these genes can be also one transcriptional unit (data not shown; Taboada et al., 2018). Therefore, the absence of OmpR in the *cas* mutants indirectly suggests that EnvZ is not produced by the polar effect of the *ompR* deletion. However, OmpR is the principal component involved in porin synthesis since the presence of the corresponding porins was reestablished in the CRISPR-*cas* deleted strain overexpessing OmpR (**Figure 4E**).

In another report, it has also been shown that a Cas protein, Cas9, negatively regulates the gene coding for a transcriptional regulator of a two-component system: *regR*. In that case, it was demonstrated that Cas9 was able to degrade the *regR* mRNA, since the *Streptococcus agalactiae* CRISPR array contains two homologous sequences to the *regR* gene (Ma et al., 2018).

Salmonella Typhi contains a Type I-E CRISPR-Cas locus, and *in vitro* experiments have demonstrated that *E. coli* components of this genetic system are able to form a complex for recognition and degradation of viral and plasmid DNA (Brouns et al., 2008; Jore et al., 2011). The data obtained in this work showed that two Cas proteins, Cas5 and Cas2, are fundamental for OmpC expression, and other four Cas proteins, Cse2, Cas6e, Cas1, and Cas3, are required for OmpF synthesis. In the case of OmpS2 expression, six Cas proteins are relevant: Cse2, Cas5, Cas6e, Cas1, Cas2, and Cas3. Therefore, it is possible that different combinations of Cas form distinct protein complexes that bind, stabilize, and positively modulate the levels of *ompR* mRNA, for differentially regulating OmpC, OmpF, or OmpS2. Another possibility for OmpR regulation is that Cse2, Cas6e, Cas1, and Cas2 RNA-nucleases cleave the mRNA of a putative *ompR* repressor. Thus, when such negative regulator would be degraded, the *ompR* gene would be able to be expressed for porin synthesis. It is also possible that the function of Cas complexes would be only to bind at DNA to fine-tune *ompR* expression at specific promoters.

Currently, experiments are being performed in our laboratory to evaluate these hypotheses and to extend these initial observations in order to define how CRISPR-Cas mediate OmpR control. It is evident that much needs to be learned about the mechanisms by which various genetic elements control the expression of the OmpR regulator and thus, the porin phenotype in *S.* Typhi.

The finding that Cas proteins are able to regulate hierarchically the global two-component regulatory systems present in different proteobacteria, suggesting that the CRISPR-Cas systems could be involved in the regulation of biological processes controlled by two-component regulators, including oxidative stress, low pH, heat shock, bacterial motility, chemotaxis, osmotic changes, resistance to bile salts, and biofilm formation (Groisman, 2016; Pruss, 2017). In this sense, OmpR regulates the expression of *hilC*, *hilD*, and *ssrAB*, the main regulators of pathogenicity islands 1 and 2 of *Salmonella* Typhimurium, and it also controls the expression of the *viaB* locus that encodes Vi polysaccharide biosynthesis genes in *S.* Typhi (Pickard et al., 1994; Lee et al., 2000; Feng et al., 2003; Cameron and Dorman, 2012). Therefore, OmpR is implicated in regulation of virulence.

In the case of the OmpC and OmpF porins, a double mutant of these genes in *S.* Typhimurium was found to be attenuated for virulence in the mouse model (Chatfield et al., 1991). In addition, it has been observed that OmpC and OmpF induced long-term antibody response with bactericidal capacity and conferred protection against challenge with *S.* Typhi (Secundino et al., 2006; Pérez-Toledo et al., 2017). Moreover, it has been demonstrated that the immunization of mice with the OmpS2 protein induced the production of specific, long-term antibody titers and conferred protection against *S.* Typhi challenge. In addition, OmpS2 is a TLR2 and TLR4 agonist. Thus, OmpS2, despite being expressed at low levels under *in vitro* culture conditions, is a potent protective immunogen with intrinsic adjuvant properties (Moreno-Eutimio et al., 2013). *Salmonella* Typhimurium mutants with deletions in the *ompS2* gene were highly attenuated for virulence in a mouse model, supporting its role in pathogenesis (Rodríguez-Morales et al., 2006).

Thus, a phenotype for the mutants in the genes coding for the *S.* Typhi Cas was explored. It was found that the *cas5* and *cas2* genes are necessary for the optimal growth of *S.* Typhi in the presence of one of the major bile salts found in the human gut, sodium deoxycholate (**Figure 5**). Most remarkably, the Δ*cas5* and Δ*cas2* mutant strains lack the OmpC porin (**Figure 2C**), which was previously shown to be necessary for growth in the presence of this bile salt (Villarreal et al., 2014).

Additionally, the CRISPR-Cas system is implicated in the control of biofilm formation in *S.* Typhi, since the absence of *cse2*, *cas5*, *cas6e*, *cas1*, *cas2*, and *cas3* genes resulted in an increase in the biosynthesis of biofilm (**Figure 6**). Interestingly, the CRISPR-Cas system has also been involved in biofilm formation in *Pseudomonas aeruginosa* (Zegans et al., 2009). These newfound roles of the *S.* Typhi CRISPR-Cas system in the resistance to sodium deoxycholate and biofilm production should contribute toward the understanding of the evolutionary conservation of this system in the *Salmonella* genus, since these biological processes are relevant for the establishment of a successful infection cycle (Gonzalez-Escobedo et al., 2011; Spector and Kenyon, 2012).

Contributions from several other research groups also support the CRISPR-Cas-outer membrane protein association. By gene neighborhood analysis, it has been found that numerous candidate CRISPR-linked genes encode integral membrane proteins in bacterial and archaeal genomes (Shmakov et al., 2018). Furthermore, activation of the CRISPR-Cas system by envelope stress has been suggested in *E. coli* (Perez-Rodriguez et al., 2011), and a role in regulating the permeability of the bacterial envelope to resist membrane damage caused by antibiotics is suggested for CRISPR-Cas in *Francisella novicida* (Sampson et al., 2014). In *Myxococcus xanthus*, the CRISPR-Cas system appears to be involved in fruiting body development and exopolysaccharide production (Viswanathan et al., 2007; Wallace et al., 2014), Moreover, recent microarray experiments performed in our laboratory demonstrated that CRISPR-Cas is able to regulate other outer membrane encoded genes besides *ompC*, *ompF*, and *ompS2* (data not shown).

Collectively, these data, together with our results suggest a previously unappreciated role for CRISPR-Cas in the formation of bacterial structures and in the maintenance of the cell envelope in different prokaryotic organisms.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

LM-A: methodology, formal analysis, investigation, writing-review, and editing. SR-G: methodology, formal analysis, and validation. JR-F: methodology and validation. AM-B, BM-M, EA-P, and AV: methodology. SE: methodology and resources. EC: writing-review and editing. IH-L: conceptualization, resources, writing-original draft preparation, visualization, supervision, project administration, and funding. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Dirección General de Asuntos del Personal Académico, DGAPA/UNAM (IN203618 and IN203621 to IH-L). SR-G was supported by a fellowship from CONACYT, Mexico (No. 887111).

ACKNOWLEDGMENTS

We would like to thank A. A. Beltrán-Luviano, M. Fernández-Mora, F. J. Santana, L. Perezgasga, I. Serrano-Fujarte, A. Tapia-Rodríguez, P. Gaytan, E. Bustos, S. Becerra, and J. Yañez from IBT-UNAM and S. Dávila from UAEM for their technical help. We also thank S. Mayo-Pérez, I. Olivar-Casique from IBT-UNAM, J. Miranda (IIBO-UNAM), and M. Dunn (CCG-UNAM) for stimulating discussions and critical reading.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.657404/full#supplementary-material

REFERENCES

Bang, I. S., Audia, J. P., Park, Y. K., and Foster, J. W. (2002). Autoinduction of the *ompR* response regulator by acid shock and control of the *Salmonella enterica* acid tolerance response. *Mol. Microbiol.* 44, 1235–1250. doi: 10.1046/j.1365-2958.2002.02937.x

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

Batchelor, E., Walthers, D., Kenney, L. J., and Goulian, M. (2005). The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins *ompF* and *ompC*. *J. Bacteriol.* 187, 5723–5731. doi: 10.1128/JB.187.16.5723-5731.2005

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

Cameron, A. D., and Dorman, C. J. (2012). A fundamental regulatory mechanism operating through OmpR and DNA topology controls expression of *Salmonella* pathogenicity islands SPI-1 and SPI-2. *PLoS Genet.* 8:e1002615. doi: 10.1371/journal.pgen.1002615

Crawford, R. W., Rosales-Reyes, R., Ramirez-Aguilar M. d. l, L., Chapa-Azuela, O., Alpuche-Aranda, C., and Gunn, J. S. (2010). Gallstones play a significant role in *Salmonella* spp. gallbladder colonization and carriage. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4353–4358. doi: 10.1073/pnas.1000862107

Chatfield, S. N., Dorman, C. J., Hayward, C., and Dougan, G. (1991). Role of *ompR*-dependent genes in *Salmonella* Typhimurium virulence: mutants deficient in both *ompC* and *ompF* are attenuated in vivo. *Infect. Immun.* 59, 449–452. doi: 10.1128/IAI.59.1.449-452.1991

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

Deiwick, J., Nikolaus, T., Erdogan, S., and Hensel, M. (1999). Environmental regulation of *Salmonella* pathogenicity island 2 gene expression. *Mol. Microbiol.* 31, 1759–1773. doi: 10.1046/j.1365-2958.1999.01312.x

De la Cruz, M. A., and Calva, E. (2010). The complexities of porin genetic regulation. *J. Mol. Microbiol. Biotechnol.* 18, 24–36. doi: 10.1159/000274309

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607. doi: 10.1038/nature09886

Dong, H., Peng, D., Jiao, X., Zhang, X., Chen, S., Lu, Y., et al. (2011). Construction and characterization of an ompR gene deletion mutant from *Salmonella enteritidis*. *Wei Sheng Wu Xue Bao* 51, 1256–1262.

Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR-Cas system. *Genes Dev.* 30, 447–459. doi: 10.1101/gad.272153.115

Encarnación, S., Guzmán, Y., Dunn, M. F., Hernández, M., del Carmen Vargas, M., and Mora, J. (2003). Proteome analysis of aerobic and fermentative metabolism in *Rhizobium etli* CE3. *Proteomics* 3, 1077–1085. doi: 10.1002/pmic.200300427

Encarnación, S., Hernández, M., Martínez-Batallar, G., Contreras, S., Vargas M. d., C., and Mora, J. (2005). Comparative proteomics using 2-D gel electrophoresis and mass spectrometry as tools to dissect stimulons and regulons in bacteria with sequenced or partially sequenced genomes. *Biol. Proced. Online* 7, 117–135. doi: 10.1251/bpo110

Feng, X., Oropeza, R., and Kenney, L. J. (2003). Dual regulation by phospho-OmpR of ssrA/B gene expression in Salmonella pathogenicity island 2. *Mol. Microbiol.* 48, 1131–1143. doi: 10.1046/j.1365-2958.2003.03502.x

Fernández-Mora, M., Puente, J. L., and Calva, E. (2004). OmpR and LeuO positively regulate the *Salmonella enterica* serovar Typhi *ompS2* porin gene. *J. Bacteriol.* 186, 2909–2920. doi: 10.1128/JB.186.10.2909-2920.2004

Ferrario, M., Ernsting, B. R., Borst, D. W., Wiese, D. E. 2nd., Blumenthal, R. M., and Matthews, R. G. (1995). The leucine-responsive regulatory protein of *Escherichia coli* negatively regulates transcription of *ompC* and *micF* and positively regulates translation of *ompF*. *J. Bacteriol.* 177, 103–113. doi: 10.1128/JB.177.1.103-113.1995

Gonzalez-Escobedo, G., Marshall, J. M., and Gunn, J. S. (2011). Chronic and acute infection of the gall bladder by *Salmonella* Typhi: understanding the carrier state. *Nat. Rev. Microbiol.* 9, 9–14. doi: 10.1038/nrmicro2490

Groisman, E. A. (2016). Feedback control of two-component regulatory systems. *Annu. Rev. Microbiol.* 70, 103–124. doi: 10.1146/annurev-micro-102215-095331

Guadarrama, C., Medrano-López, A., Oropeza, R., Hernández-Lucas, I., and Calva, E. (2014). The *Salmonella enterica* serovar Typhi LeuO global regulator forms tetramers: residues involved in oligomerization, DNA binding, and transcriptional regulation. *J. Bacteriol.* 196, 2143–2154. doi: 10.1128/JB.01484-14

Hernández-Lucas, I., Gallego-Hernández, A. L., Encarnación, S., Fernández-Mora, M., Martínez-Batallar, A. G., Salgado, H., et al. (2008). The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 190, 1658–1670. doi: 10.1128/JB.01649-07

Hurkman, W. J., and Tanaka, C. K. (1986). Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol.* 81, 802–806. doi: 10.1104/pp.81.3.802

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

Kawaji, H., Mizuno, T., and Mizushima, S. (1979). Influence of molecular size and osmolarity of sugars and dextrans on the synthesis of outer membrane proteins O-8 and O-9 of *Escherichia coli* K-12. *J. Bacteriol.* 140, 843–847. doi: 10.1128/JB.140.3.843-847.1979

Lee, A. K., Detweiler, C. S., and Falkow, S. (2000). OmpR regulates the two-component system SsrA-ssrB in *Salmonella* pathogenicity island 2. *J. Bacteriol.* 182, 771–781. doi: 10.1128/JB.182.3.771-781.2000

Lee, Y. H., Kim, B. H., Kim, J. H., Yoon, W. S., Bang, S. H., and Park, Y. K. (2007). CadC has a global translational effect during acid adaptation in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* 189, 2417–2425. doi: 10.1128/JB.01277-06

Lu, Y., Chen, S., Dong, H., Sun, H., Peng, D., and Liu, X. (2012). Identification of genes responsible for biofilm formation or virulence in *Salmonella enterica* serovar pullorum. *Avian Dis.* 56, 134–143. doi: 10.1637/9806-052411-Reg.1

Ma, K., Cao, Q., Luo, S., Wang, Z., Liu, G., Lu, C., et al. (2018). *cas9* enhances bacterial virulence by repressing the *regR* transcriptional regulator in *Streptococcus agalactiae*. *Infect. Immun.* 86, e00552–e005517. doi: 10.1128/IAI.00552-17

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). 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

Martínez-Laguna, Y., Calva, E., and Puente, J. L. (1999). Autoactivation and environmental regulation of *bfpT* expression, the gene coding for the transcriptional activator of *bfpA* in enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 33, 153–166. doi: 10.1046/j.1365-2958.1999.01460.x

Medina-Aparicio, L., Dávila, S., Rebollar-Flores, J. E., Calva, E., and Hernández-Lucas, I. (2018). The CRISPR-Cas system in *Enterobacteriaceae*. *Pathog. Dis.* 76. doi: 10.1093/femspd/fty002

Medina-Aparicio, L., Rebollar-Flores, J. E., Beltrán-Luviano, A. A., Vázquez, A., Gutíerrez-Ríos, R. M., Olvera, L., et al. (2017). CRISPR-Cas system presents multiple transcriptional units including antisense RNAs that are expressed in minimal medium and upregulated by pH in *Salmonella enterica* serovar Typhi. *Microbiology* 163, 253–265. doi: 10.1099/mic.0.000414

Medina-Aparicio, L., Rebollar-Flores, J. E., Gallego-Hernández, A. L., Vázquez, A., Olvera, L., Gutíerrez-Ríos, R. M., et al. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 193, 2396–2407. doi: 10.1128/JB.01480-10

Meng, J., Bai, J., Xu, J., Huang, C., and Chen, J. (2019). Differential regulation of physiological activities by RcsB and OmpR in *Yersinia enterocolitica*. *FEMS Microbiol. Lett.* 366:fnz210. doi: 10.1093/femsle/fnz210

Moreno-Eutimio, M. A., Tenorio-Calvo, A., Pastelin-Palacios, R., Pérez-Shibayama, C., Gil-Cruz, C., López-Santiago, R., et al. (2013). *Salmonella* Typhi OmpS1 and OmpS2 porins are potent protective immunogens with adjuvant properties. *Immunology* 139, 459–471. doi: 10.1111/imm.12093

Oropeza, R., Salgado-Bravo, R., and Calva, E. (2015). Deletion analysis of RcsC reveals a novel signalling pathway controlling poly-N-acetylglucosamine synthesis and biofilm formation in *Escherichia coli*. *Microbiology* 161, 903–913. doi: 10.1099/mic.0.000050

O'Toole, G. A. (2011). Microtiter dish biofilm formation assay. *J. Vis. Exp.* 47:2437. doi: 10.3791/2437

Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K. H., Bernardis, S., Ke, A., et al. (2011). Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.* 79, 584–599. doi: 10.1111/j.1365-2958.2010.07482.x

Pérez-Toledo, M., Valero-Pacheco, N., Pastelin-Palacios, R., Gil-Cruz, C., Pérez-Shibayama, C., Moreno-Eutimio, M. A., et al. (2017). *Salmonella* Typhi porins OmpC and OmpF are potent adjuvants for T-dependent and T-independent antigens. *Front. Immunol.* 8:230. doi: 10.3389/fimmu.2017.00230

Pickard, D., Li, J., Roberts, M., Maskell, D., Hone, D., Levine, M., et al. (1994). Characterization of defined ompR mutants of *Salmonella* Typhi: ompR is involved in the regulation of Vi polysaccharide expression. *Infect. Immun.* 62, 3984–3993. doi: 10.1128/IAI.62.9.3984-3993.1994

Pruss, B. M. (2017). Involvement of two-component signaling on bacterial motility and biofilm development. *J. Bacteriol.* 199, e00259–e002517. doi: 10.1128/JB.00259-17

Puente, J. L., Flores, V., Fernández, M., Fuchs, Y., and Calva, E. (1987). Isolation of an *ompC*-like outer membrane protein gene from *Salmonella* Typhi. *Gene* 61, 75–83. doi: 10.1016/0378-1119(87)90366-0

Puente, J. L., Juárez, D., Bobadilla, M., Arias, C. F., and Calva, E. (1995). The *Salmonella ompC* gene: structure and use as a carrier for heterologous sequences. *Gene* 156, 1–9. doi: 10.1016/0378-1119(94)00883-T

Rodríguez-Morales, O., Fernández-Mora, M., Hernández-Lucas, I., Vázquez, A., Puente, J. L., and Calva, E. (2006). *Salmonella enterica* serovar Typhimurium *ompS1* and *ompS2* mutants are attenuated for virulence in mice. *Infect. Immun.* 74, 1398–1402. doi: 10.1128/IAI.74.2.1398-1402.2006

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

Samanta, P., Clark, E. R., Knutson, K., Horne, S. M., and Pruss, B. M. (2013). OmpR and RcsB abolish temporal and spatial changes in expression of *flhD* in *Escherichia coli* biofilm. *BMC Microbiol.* 13:182. doi: 10.1186/1471-2180-13-182

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sampson, T. R., Napier, B. A., Schroeder, M. R., Louwen, R., Zhao, J., Chin, C. Y., et al. (2014). A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. *Proc. Natl. Acad. Sci. U. S. A.* 111, 11163–11168. doi: 10.1073/pnas.1323025111

Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y. L., and Weiss, D. S. (2013). A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497, 254–257. doi: 10.1038/nature12048

Secundino, I., López-Macias, C., Cervantes-Barragan, L., Gil-Cruz, C., Ríos-Sarabia, N., Pastelin-Palacios, R., et al. (2006). *Salmonella* porins induce a sustained, lifelong specific bactericidal antibody memory response. *Immunology* 117, 59–70. doi: 10.1111/j.1365-2567.2005.02263.x

Shariat, N., Timme, R. E., Pettengill, J. B., Barrangou, R., and Dudley, E. G. (2015). Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology* 161, 374–386. doi: 10.1099/mic.0.000005

Shin, S., and Park, C. (1995). Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. *J. Bacteriol.* 177, 4696–4702. doi: 10.1128/JB.177.16.4696-4702.1995

Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Severinov, K. V., and Koonin, E. V. (2018). Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl. Acad. Sci. U. S. A.* 115, E5307–E5316. doi: 10.1073/pnas.1803440115

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

Spector, M. P., and Kenyon, W. J. (2012). Resistance and survival strategies of *Salmonella enterica* to environmental stresses. *Food Res. Int.* 45, 455–481. doi: 10.1016/j.foodres.2011.06.056

Taboada, B., Estrada, K., Ciria, R., and Merino, E. (2018). Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics* 34, 4118–4120. doi: 10.1093/bioinformatics/bty496

Touchon, M., and Rocha, E. P. (2010). The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS One* 5:e11126. doi: 10.1371/journal.pone.0011126

Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M., and Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new *ompR* allele that increases curli expression. *J. Bacteriol.* 180, 2442–2449. doi: 10.1128/JB.180.9.2442-2449.1998

Villarreal, J. M., Becerra-Lobato, N., Rebollar-Flores, J. E., Medina-Aparicio, L., Carbajal-Gómez, E., Zavala-García, M. L., et al. (2014). The *Salmonella*

enterica serovar Typhi *ltrR-ompR-ompC-ompF* genes are involved in resistance to the bile salt sodium deoxycholate and in bacterial transformation. *Mol. Microbiol.* 92, 1005–1024. doi: 10.1111/mmi.12610

Viswanathan, P., Murphy, K., Julien, B., Garza, A. G., and Kroos, L. (2007). Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.* 189, 3738–3750. doi: 10.1128/JB.00187-07

Wallace, R. A., Black, W. P., Yang, X., and Yang, Z. (2014). A CRISPR with roles in *Myxococcus xanthus* development and exopolysaccharide production. *J. Bacteriol.* 196, 4036–4043. doi: 10.1128/JB.02035-14

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

Yoshida, T., Qin, L., Egger, L. A., and Inouye, M. (2006). Transcription regulation of *ompF* and *ompC* by a single transcription factor, OmpR. *J. Biol. Chem.* 281, 17114–17123. doi: 10.1074/jbc.M602112200

Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., and O'Toole, G. A. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191, 210–219. doi: 10.1128/JB.00797-08

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.

frontiers in Microbiology

MINI REVIEW
published: 31 March 2021
doi: 10.3389/fmicb.2021.635227

Gene Silencing Through CRISPR Interference in Bacteria: Current Advances and Future Prospects

Riyu Zhang[1†], Wensheng Xu[2†], Shuai Shao[1] and Qiyao Wang[1,3,4]*

[1] State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, China, [2] Laboratory of Agricultural Product Detection and Control of Spoilage Organisms and Pesticide Residue, Faculty of Food Science and Engineering, Beijing University of Agriculture, Beijing, China, [3] Shanghai Collaborative Innovation Center for Biomanufacturing Technology, Shanghai, China, [4] Shanghai Engineering Research Center of Maricultured Animal Vaccines, Shanghai, China

OPEN ACCESS

Edited by:
Aixin Yan,
The University of Hong Kong, Hong Kong

Reviewed by:
Zhi-Qiang Xiong, University of Shanghai for Science and Technology, China
Yang Gu, Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Shuai Shao
shaoscott@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *01 December 2020*
Accepted: *24 February 2021*
Published: *31 March 2021*

Citation:
Zhang R, Xu W, Shao S and Wang Q (2021) Gene Silencing Through CRISPR Interference in Bacteria: Current Advances and Future Prospects. Front. Microbiol. 12:635227. doi: 10.3389/fmicb.2021.635227

Functional genetic screening is an important method that has been widely used to explore the biological processes and functional annotation of genetic elements. CRISPR/Cas (Clustered regularly interspaced short palindromic repeat sequences/CRISPR-associated protein) is the newest tool in the geneticist's toolbox, allowing researchers to edit a genome with unprecedented ease, accuracy, and high-throughput. Most recently, CRISPR interference (CRISPRi) has been developed as an emerging technology that exploits the catalytically inactive Cas9 (dCas9) and single-guide RNA (sgRNA) to repress sequence-specific genes. In this review, we summarized the characteristics of the CRISPRi system, such as programmable, highly efficient, and specific. Moreover, we demonstrated its applications in functional genetic screening and highlighted its potential to dissect the underlying mechanism of pathogenesis. The recent development of the CRISPRi system will provide a high-throughput, practical, and efficient tool for the discovery of functionally important genes in bacteria.

Keywords: CRISPR/Cas, gene silencing, essential genes, Tn-seq, CRISPRi screen

INTRODUCTION

Genome editing is a robust technology of modifying genome with a high efficiency emerging in recent years, which has a growing and profound influence on bioscience, biotechnology, and bio-industry. ZFN and TALEN, the primary generation genome editing technologies, are protein-guided and need protein engineering, which is time-consuming and not easy to operate (Kim et al., 1996; Cathomen and Joung, 2008; Wood et al., 2011). CRISPR/Cas-based genome editing hereby came into being and could theoretically edit the genome of any organism. So far, the powerful technology has brought about a revolution in biology due to its significantly simplified construction process.

In general, essential genes are hard to be probed because their knock-out is lethal to the organism. Gene silencing technology such as RNA interference (RNAi) is capable of inhibiting the expression of genes and, hence, is applied to investigate the function of essential genes (Agrawal et al., 2003). Since RNAi is mainly utilized in eukaryotes, a silencing tool is also required for dissecting the essential genes in the prokaryote system. CRISPRi has been developed from

the CRISPR/Cas-based genome editing to fill in the blanks. Furthermore, coupling with high-throughput sequencing, it has emerged as a potential and promising strategy to perform functional genomics research in bacteria. Here, this review gives a brief introduction to the CRISPRi system, the underlying mechanism and properties, and highlights its application as a high-throughput screening tool in gene functional analysis.

THE MECHANISM OF THE CRISPR/CAS-BASED GENOME EDITING AND INTERFERENCE SYSTEM

CRISPR/Cas-based genome editing is a newly developed RNA-guided genome editing system. CRISPR is a series of clustered DNA sequences including repeats and spacers and Cas are CRISPR-associated proteins (Jansen et al., 2002). The CRISPRs are observed in nearly 90% of genomes of the sequenced archaea and nearly 40% of genomes of the sequenced bacteria (Sorek et al., 2008). They can be divided into two classes based on the number of Cas proteins interfering with an invading DNA (Makarova et al., 2015). Class 1 systems include type I, III, and IV, which have multi-subunit effector complexes and, hence, are not suitable to be applied to genome editing. Instead, Class 2 systems consist of type II, V, and VI, which only possess a single effector protein. Moreover, type II is the simplest CRISPR/Cas systems and can achieve interference with an invading DNA only by a single multi-functional effector Cas protein (Makarova et al., 2011).

As the adaptive immune systems of prokaryote, CRISPR/Cas systems can recognize and cleave foreign nucleic acids (Barrangou et al., 2007; Brouns et al., 2008). With the help of a chimeric single guide RNA (sgRNA), the Cas protein is targeted to a specific DNA sequence and then triggers a double-strand break (DSB) at the chromosomal DNA (Deltcheva et al., 2011). The recognized short DNA sequence is called the protospacer adjacent motif (PAM) and different Cas proteins can recognize different PAM sequences. For instance, the recognized PAM of SpCas9 protein derived from *Streptococcus pyogenes* is NGG whereas the recognized PAM of StCas9 from *Streptococcus thermophilus* is NGGNG (Cho et al., 2013; Karvelis et al., 2013). Coupled with an available and editing template DNA, such DSBs could be repaired through homologous recombination (HR) to introduce precise genome editing. Instead, such DSBs could also be repaired by Non-homologous end joining (NHEJ), which would produce small insertion and deletion mutations to abolish or disrupt the function of the target gene.

As one of the most commonly used Cas proteins, the Cas9 protein can cleavage the invading DNA because it possesses RuvC and HNH nuclease domains which can cleave the non-complementary strand and the complementary strand, respectively (Jinek et al., 2012). The catalytic domains of Cas9 are mutated to generate the inactive dCas9 (nuclease – dead mutants of Cas9) lacking the endonuclease activity but instead, it still can be in conjunction with the sgRNA (Jinek et al., 2012). Consequently, the dCas9-sgRNA complex specifically binds to the target gene at the promoter or coding sequence and acts as a roadblock to the elongating RNA polymerase, hence, aborting transcription initiation, or elongation. The function of dCas9 was confirmed by the native elongating transcript sequencing (NET-seq) experiment (Qi et al., 2013) and **Figure 1** demonstrated the RNAP is blocked by the dCas9-sgRNA complexes.

To achieve a CRISPR/Cas-based interference, scientists have already developed multiple strategies such as the plasmid-based system and the chromosomally integrated system. The plasmid-based system consists of the single-plasmid system and the dual-plasmid system. The single-plasmid system employs a composite plasmid harboring dCas9 and sgRNA together, while the major limitation is the cloning efficiency due to the relatively large plasmid size. The construction of the dual-plasmid system is simplified, in which dCas9 and sgRNA are carried by two independent small plasmids. However, the plasmid incompatibility and stability have to be taken into consideration before applying it. Both of the plasmid-based systems have no need to integrate elements into the genome of bacteria, avoiding the unexpected consequences of the change of genome, and have been extensively applied to silence single gene or multiple genes in bacteria (see section "Application of CRISPRi in Bacteria"). Alternatively, the chromosomally integrated system was developed and the dCas9 is integrated into a neutral site of the bacterial genome. The sgRNA exists in a small plasmid similar to that of the plasmid-based system and the establishment of sgRNA assessment algorithm enables the design of a high-saturated sgRNA plasmid library. With the decreasing cost of DNA synthesis, it is feasible to synthesize large sgRNA libraries, leveraging the chromosomally integrated system for almost all the high-throughput CRISPRi screening in bacteria (see section "Application of CRISPRi in High-Throughput Screen").

PROPERTIES OF CRISPRi

The CRISPRi technique is originated from CRISPR and, thus, possesses many properties as same as what CRISPR owned. They are both programmable, highly efficient, and specific but also face the troubles of off-target and toxicity. What is more, there are some distinct properties from the CRISPR/Cas-based genome editing including both the merits and demerits (**Table 1**).

Many advantages emerge in CRISPRi. First of all, CRISPRi could simultaneously regulate the expression of multiple genes, expanding the breadth of this application. Qi et al. (2013) found that CRISPRi could be applied to regulate multiple genes independently without crosstalk. Kim et al. (2017) used the tunable CRISPRi system to repress the expression of multiple genes and achieved the increment of n-butanol yield and productivity in recombinant *Escherichia coli*. Lv et al. (2015) utilized CRISPRi to manipulate the expression of multiple essential genes involved in 4HB synthesis and then regulate P(3HB-co-4HB) composition.

Not only the expression of genes, but also the degree of gene repression could be controlled. Tuning gene repression is helpful because some genes are extremely sensitive to knockdown and many genes of interest are expected to be expressed under tight control. This can be fulfilled by titrating the concentration of dCas9 or sgRNA from an inducible promoter, which is easy and

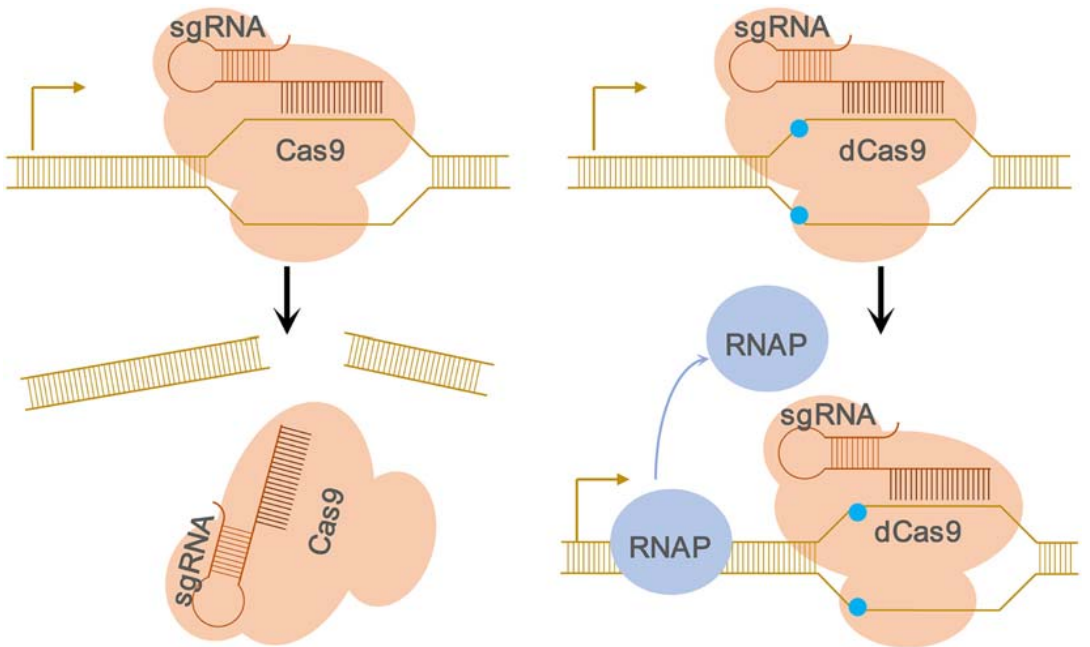


FIGURE 1 | Schematic of the CRISPRi approach. **Left**, Cartoon representation of the CRISPR-mediated regulation of gene expression. The wild-type Cas9 protein binds to the sgRNA and forms a protein–RNA complex. Once Cas9–sgRNA complex binds to specific DNA target adjacent to PAM, it leads to the cleavage of the target DNA due to the nuclease activity of the Cas9 protein. **Right**, Cartoon depicting the CRISPRi-mediated interference of gene expression caused by nuclease-deficient dCas9. The nuclease-deficient dCas9 contains two substitutions in the nuclease domains (D10A and H840A, blue dots), and thus lose the endonuclease activity. If the target DNA sequence locates inside an open reading frame, the dCas9–sgRNA–DNA complex will block the movement of RNAP and subsequent transcription elongation, resulting in transcription inhibition of the target gene.

straightforward. Li et al. (2016) modulated the expression levels of target genes via controlling the expression of dCas9 under the control of P_{BAD} promoter, resulting in over two orders of magnitude dynamic range. Fontana et al. (2018) achieved a broad range of titration of the CRISPRi repression by changing the level of gRNA from the P_{tet} promoter in *E. coli*. Partial repression of genes could also be realized by introducing mismatches between the target and gRNA. The method is suitable when the bacteria is sensitive to the level of dCas9 required for maintaining at a low concentration. Bikard et al. (2013) took advantage of the mismatches between the crRNA and target DNAs to modulate their repression level in *E. coli*.

In addition, CRISPRi-based knockdown is inducible and reversible, which enables the temporal and dynamic regulation of interested genes. When dCas9 is under the tight control of the anhydrotetracycline-inducible (aTc-inducible) promoter, the knockdown could be either be induced by aTc or reversed by removing the inducer from the culture (Qi et al., 2013). Likewise, the arabinose could activate the expression of dCas9 from the P_{BAD} promoter to induce a CRISPRi-based knockdown and this silencing is reversible once the inducer is washed away from the media (Li et al., 2016). Wang T. et al. (2019) found that genes of *Yersinia pestis* silenced by CRISPRi might restore expression by washing away the inducer in an actively replicating bacteria.

On the other hand, there are many disadvantages in CRISPRi including the bad-seed effect, polar effect and reverse polar effect, toxicity, and off-target. Bad-seed effect was defined as sgRNAs with specific 5-nt seed sequences that can produce strong fitness defects regardless of the other 15-nt of the guide sequence (Cui et al., 2018). Similarly, polar effect and reverse polar effect could disturb the result of the CRISPRi screen. Peters et al. (2016) reported the polar effect when sgRNA blocking a gene will repress the expression of all downstream genes in an operon. It is expected because once blocked by dCas9, RNAP is hard to go forward and trigger the transcription of downstream genes (Peters et al., 2016). In addition, when the dCas9-sgRNA complexes were guided to non-essential genes located upstream of the essential genes in operons, it exhibited a strong impairment on cell fitness (Cui et al., 2018). It is not surprising that disrupting the transcription of an upstream gene will cause a depletion of the cotranscribed downstream genes as well because they are often carried on a single transcript. Intriguingly, the reverse polar effect was also observed when the dCas9-sgRNA complexes might silence the upstream of targeted genes (Peters et al., 2016; Cui et al., 2018). The reverse polar effect could be explained due to the destabilization of the interrupted transcript.

Like Cas9, the degree of toxicity of dCas9 has been demonstrated in many bacteria. Lee et al. (2019) found that dCas9 can lead to a longer lag phase of *Vibrio natriegens*, which indicated the marginal toxicity of dCas9. The high-level dCas9 severely decreased the growth rate of *E. coli* and changed the cell shape to an abnormal filamentous morphology (Cho et al., 2018). High-level dCas9 up-regulated the genes associated with cell division and down-regulated the genes encoding proteins located in the

cell membrane. The dCas9 directly bound upstream of 37 genes without sgRNA including *fimA* encoding bacterial fimbriae. A high concentration of $dCas9_{Spy}$ was lethal to *Mycobacterium tuberculosis* without a target sgRNA (Rock et al., 2017). Instead, Zhang and Voigt constructed a non-toxic version of dCas9 (dCas9*_PhlF) to avoid off-target effects, that binds to DNA through PhlF instead of dCas9 (Zhang and Voigt, 2018). Hence, it is worth noting that the effect on the growth of recipient strains by dCas9 should be tested before applying CRISPRi.

Furthermore, off-target effects appeared in CRISRPi. The sgRNAs with 9-nt of identity in the seed sequence can produce off-target effects in *E. coli* MG1655 (Cui et al., 2018). In general, off-targeting is rarely encountered in bacteria, in part because the relatively small genome size of bacteria limits the potential for sites with only one or two mismatches similar to the target sequence (Bikard et al., 2013; Rock et al., 2017). Indeed, we could not rule out the possibility of off-target effects, hence, target sites should be cautiously determined to ensure that they are unique and not highly similar to other sites in the genome.

COMPARISON BETWEEN THE CRISPRi AND sRNA-MEDIATED GENE SILENCING

Before the emergence of CRISPRi, the gene silencing by RNA interference (RNAi) with small RNA (sRNA), a series of RNA whose length is within 50–200 nt such as miRNA and siRNA, have been developed as a powerful tool of gene silencing in eukaryotes. As shown in **Table 2**, there are many common characteristics between the sRNA-mediated gene silencing and CRISPRi technique including programmable, off-target, simultaneous inhibition of the expression of multiple genes, etc. However, the CRISPRi technique has its distinct properties. First, the inhibition of the initiation and elongation of RNAP by the dCas9-sgRNA complex is at a transcription level, while the block of the initiation of the ribosome by sRNA is at a post-transcription level. Second, the objects and sites of the target are different. The target of sRNA-mediated gene silencing is commonly the 5′ UTR of mRNA while CRISPRi will target the promoter or the ORF of interesting genes (Na et al., 2013), which enables a more stable and efficient interference. Lastly, the sRNA-mediated gene silencing mainly needs the help of other chaperon proteins. For instance, the RNA chaperon Hfq is required for base pairing and thus stabilize the interaction between sRNA and its target mRNA (Beisel and Storz, 2010). Whereas, CRISPRi functions via dCas9 protein acting as a roadblock (Man et al., 2011). Therefore, CRISPRi is a more robust and widely used tool in gene silencing.

COMPARISON BETWEEN CRISPRi SCREEN AND Tn-seq

Genome-wide screening could associate genes with phenotypes at a large scale in bacteria. Transposon sequencing was applied extensively in functional genomics research (van Opijnen et al., 2009). A saturated transposon insertion library is cultured in competitive conditions and the fitness for each mutant can be determined through NGS analysis. However, there are some certain technical limitations that it is impossible to investigate essential genes because the transposons inserted strains might exhibit growth deficiency and that not all Tn-insertions result in gene inactivation. Accordingly, the Tn-seq screen requires large libraries to fully cover the genome and eventually (Yang et al., 2017), this will lead to a large amount of nonsense mutations. The complexity of mutant libraries may cause a bottleneck effect during the following screens.

CRISPR interference screen could facilitate functional analysis of essential genes because it can probe both non-essential and essential genes. Moreover, the CRISPRi screen is not only applicable to genome-wide but also a tiling library for specific genes of interest, which is cost-saving. The genome-wide library or tiling library could be selected when designing sgRNA. Hence, the CRISPRi screen library is more flexible than the Tn-seq due to its adjustable library size (**Table 3**). In addition, the CRISPRi screen is more suitable for mapping phenotypes to short genes than the Tn-seq as the latter may produce poor statistical robustness when short genes such as non-coding RNAs are investigated (Wang et al., 2018). However, the off-target effects have hindered the application of the CRISPRi screen. In this view, the Tn-seq has an obvious advantage over the CRISPRi screen in terms of the accuracy, therefore, the combination of the CRISPRi screen and Tn-seq will provide a reasonable and effective strategy in functional genomics research.

APPLICATION OF CRISPRi IN BACTERIA

CRISPR interference could be applied to perform a functional analysis of specific genes in pathogens. Scientists established a CRISPRi-based knockdown system to achieve the attenuation of virulence in the animal model when virulence genes were

TABLE 1 | The advantages and disadvantages of CRISPRi.

Advantages	Disadvantages
Tunable knockdown	Off-target
Inducible knockdown	Bad-seed effect
Reversible knockdown	Polar effect and reverse polar effect
Controlling multiple genes	Toxicity of dCas9

TABLE 2 | Comparison of sRNA-mediated gene silencing and CRISPRi.

Characteristic	sRNA-mediated gene silencing	CRISPR interference
Silence type	post-transcriptional inhibition	transcription inhibition
Target	5′ UTR of mRNA	promoter or the ORF of interested genes
Required tools	sRNA and chaperone protein (Hfq, etc.)	sgRNA and dCas9
Common characteristics	silencing through base pairing of RNA and target, silence simultaneously multiple genes, programmable, off-target	

TABLE 3 | Comparison of Tn-seq and CRISPRi screen.

Characteristic	Tn-seq	CRISPRi screen
Mutation type	insertion mutation	knock-down
Required tools	transposon	sgRNA, dCas9
Gene selection	random insertion	random design or specific design
Gene type	non-essential gene	non-essential gene or essential gene
Library size	genome-wide library	genome-wide library or tiling library
Advantages	suitable for operon genes	adjustable library size, suitable for essential gene, and short genes
Disadvantages	unsuitable for essential gene and short genes, transposon insertion prefers TA site, which results in insertion sites uneven distribution.	off-target, polar and reverse polar effect, bad-seed effect, dCas9-specific toxicity

silenced, which lay the foundation for probing the virulence-associated essential genes of the unknown function in pathogens. The Mobile-CRISPRi system was established including modular and transferable components that can be integrated into the genomes of diverse bacteria to expand the range of the CRISPRi systems within bacteria (Peters et al., 2019). Mobile-CRISPRi was used to control the expression of conditionally essential (CE) virulence genes in a murine model of pneumonia with the purpose of dissecting the function of CE genes. Based on this analysis, the gene *exsA*, a CE gene encoding the type III secretion system activator, was identified to repress and inhibit the secretion of effectors and attenuate virulence in mice (Qu et al., 2019). Wang T. et al. (2019) introduced an optimized CRISPRi system into *Yersinia pestis* and thereby repressed virulence-associated genes *yscB* or *ail*, resulting in the virulence attenuation in HeLa cells and mice, in line with the previously reported phenotypes caused by *yscB* and *ail* knockout.

By introducing CRISPRi into the bacterial pathogen, the essential genes intimately tied to viability can be characterized, which provided several novel targets for vaccine and antibiotic development. Caro et al. (2019) focused on essential genes involved in viability and virulence in *Vibrio cholerae* and identified that the reduced expression of the lipoprotein transport (Lol) system rendered cells prone to plasmolysis and resulted in dynamic membrane rearrangements and extrusion of mega outer membrane vesicles, which thus provided a novel drug target. In *Streptococcus pneumoniae* serotype 2 strain D39, the genes *murT* and *gatD* were determined as the essential genes for peptidoglycan synthesis (Liu et al., 2017). Also, *tarP* and *tarQ* involving in the polymerization of teichoic acid precursors were also identified (Liu et al., 2017), which would contribute to the development of novel vaccines and antibiotics.

Besides, CRISPRi could be applied to the biosynthesis of commodity chemicals. Scientists utilized CRISPRi to increase the production of chemicals in bacteria. CRISPRi was used to silence the genes on the branch pathways of surfactin synthesis and thereby enhance the amino acid precursor supply in order to increase the production of surfactin in *Bacillus subtilis* (Wang C. et al., 2019). Improved production of anthocyanin peonidin 3-O-glucoside (P3G) was realized through repressing the transcriptional repressor MetJ to downregulate the methionine biosynthetic pathway in *E. coli* (Cress et al., 2017). In *Corynebacterium glutamicum*, the shikimic acid yield was increased by altering the expression of related genes (Zhang et al., 2016).

Apart from the biosynthesis of natural chemicals, the chemical composition can be modified by CRISPRi. For example, polyhydroxyalkanoates (PHA), a family of biodegradable and biocompatible polyesters consisting of poly (3-hydroxybutyrate-co-4-hydroxybutyrate) [P(3HB-co-4HB)], possesses similar properties with traditional plastics with its physical nature depending on the ratio of 3HB and 4HB. In order to meet the demand of diverse industrial applications, the expression

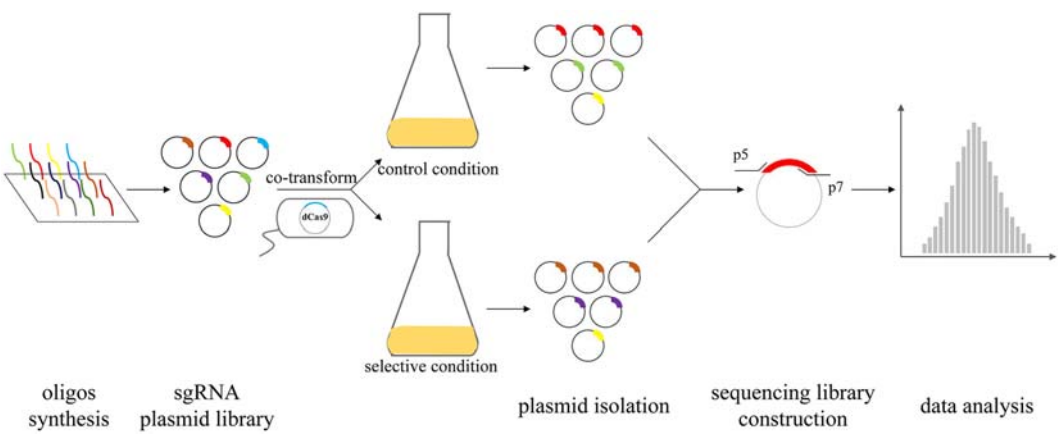


FIGURE 2 | The CRISPRi screening procedure in bacteria. Firstly, oligos designed on a large scale by the software are synthesized through a chip and then, the sgRNA plasmids library is constructed. Mix sgRNA plasmids are co-transformed into recipient strains in which dCas9 is integrated into genomic DNA. Competitive growth culture is performed in selective condition and control conditions. The mixed plasmids are isolated from pooled colonized strain and subsequently used for high-throughput sequencing. Finally, data analysis of the relative fitness contribution of each gene is performed.

of multiple essential genes involved in 4HB synthesis was manipulated resulting in the modification of P(3HB-co-4HB) composition (Lv et al., 2015). As another example, the recombinant *E. coli* harboring *phaCAB* operon was widely applied to produce Polyhydroxybutyrate (PHB). The activity of PhaC is direct to PHB accumulation but reverse to PHB molecular weight. Li et al. (2017) precisely controlled the expression of *phaC* by targeting diverse sites by sgRNA with the aim of modulating the balance between PHB accumulation and PHB molecular weight.

APPLICATION OF CRISPRi IN HIGH-THROUGHPUT SCREEN

Combined with high-throughput screen work, CRISPRi has been performed extensively to investigate the phenotypes of essential genes involved mainly in cell morphology and growth in prokaryotes. Peters et al. (2016) used the CRISPRi screening to discover essential genes intimately tied to cell morphology in the Gram-positive model bacterium *Bacillus subtilis*. Lee et al. (2019) identified a minimal set of genes required for the rapid growth of the fast-growing bacterium *Vibrio natriegens* contributing to further research and engineering of *Vibrio natriegens*. Rousset et al. (2018) identified *E. coli* genes required by phages λ, T4, and 186 for the production of functional progeny, which thus provided novel insights into the design of improved phage therapies. A thorough CRISPRi screening in *Synechocystis* sp. PCC 6803 lead to the identification of *gltA* and *pcnB* facilitating the productivity of L-lactate and *bcp2*, the L-lactate tolerance related gene (Yao et al., 2020). A pooled CRISPRi screen facilitated the discovery of growth switches *sibB*/*ibsB*, which can be applied to decoupling cell growth and protein production in *E. coli* (Li et al., 2020).

Furthermore, the CRISPRi screen could be used to reveal the properties and design rules of itself. Coupling the CRISPRi screen with machine learning, the bad-seed effect was found and furthermore, it could be alleviated by the reduced dCas9 concentration (Cui et al., 2018). Wang et al. (2018) reported that sgRNAs targeting to the first 5% of the ORF near the start codon increased the efficiency of silencing in *E. coli* MG1655 and they also defined 10 sgRNAs/gene as the minimal sufficient number for reliable hit-gene calling. Calvo-Villamanan et al. (2020) provided a novel model to predict on-target activity for dCas9 based on the target sequence in *E. coli*, especially in bases surrounding the PAM sequence when dCas9 binds to the coding strand. The result indicated that the silencing activity of dCas9 was not only determined by sgRNA but also by the target sequence (Calvo-Villamanan et al., 2020).

FUTURE PERSPECTIVES

Over the last few years, high-throughput CRISPRi screens have been performed with a veritable explosion of high-throughput, high-dimensional of essential genes, and conditional essential genes. These studies have revealed numerous new biology, not only novel gene functions but also novel connections within gene networks.

As a promising tool in bacterial genome engineering, CRISPRi screens have become increasingly common in diverse bacteria. However, the main challenge for CRISPRi application in bacteria is the off-target effect. To overcome this hurdle, the engineering of dCas9 variants is on-going. Recently, an expanded PAM SpCas9 variant (xCas9) was established and has a much greater DNA specificity than most commonly used SpCas9, resulting in substantially lower off-target effects at genome-wide targeted sites (Hu et al., 2018). On the other hand, optimizing sgRNA design is thought to be a useful strategy to minimize off-target effects. In *E. coli*, a high-density and comprehensive sgRNA on-target activity map was constructed and then used to guide the optimization of sgRNA on-target activity prediction algorithm, aiming to accurately predict highly effective sgRNAs (Guo et al., 2018).

A powerful tool for silencing genes is required for dissecting the underlying mechanism of pathogenesis. Compared to transposon-based approaches, CRISPRi libraries are more compact, which can enable proper use in situations where Tn-seq would be bottlenecked with similar genomic coverage. The sgRNA plasmid design, which contains the 20-bp target sequence, facilitates an easy and efficient cloning that is readily scalable for the construction of sgRNA libraries for genome-wide gene targeting. A concise pooled CRISPR interference system was recently built for high-throughput quantitative genetic interaction screening on a genome-wide scale for the important human pathogen *S. pneumoniae* (Liu et al., 2021). Several *S. pneumoniae* genes were identified as essential in a laboratory medium whereas exhibited neutral in the host. This gives us a hint that the role of essential genes of pathogens might be overlooked during the infection due to conventional mutagenesis. Based on this idea, high-throughput *in vivo* evaluation of the fitness cost of genes by CRISPRi screen can broaden our horizons of pathogenesis research (**Figure 2**). CRISPRi screen can be used to probe essential genes that are difficultly characterized by Tn-seq, thus, is a complement to Tn-seq in high-throughput methods to facilitate the mapping of genotype-phenotype associations of core and more essential genes. The CRISPRi screen has the prospect of becoming a powerful tool screening vaccine and drug targets because many essential genes that are potential targets previously rarely researched. The knowledge of conditional essential genes acquired from CRISPRi screen will provide new insights and expand our current understanding of the functional genomics in prokaryotes.

AUTHOR CONTRIBUTIONS

SS and QW designed the mini review. All authors wrote, revised, and approved the manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (32002436 to SS and 31772891

to QW), Shanghai Sailing Program (20YF1411100), China Postdoctoral Science Foundation (2020M671034), and Beijing Municipal Commission of Education (KM20161002 0016).

ACKNOWLEDGMENTS

We are grateful to Jianchang Huang, Yue Peng, and Haoyuan Ding for their helpful discussions.

REFERENCES

Agrawal, N., Dasaradhi, P. V., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., and Mukherjee, S. K. (2003). RNA interference: biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* 67, 657–685. doi: 10.1128/mmbr.67.4.657-685.2003

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

Beisel, C. L., and Storz, G. (2010). Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol. Rev.* 34, 866–882. doi: 10.1111/j.1574-6976.2010.00241.x

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41, 7429–7437. doi: 10.1093/nar/gkt520

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

Calvo-Villamanan, A., Ng, J. W., Planel, R., Menager, H., Chen, A., Cui, L., et al. (2020). On-target activity predictions enable improved CRISPR-dCas9 screens in bacteria. *Nucleic Acids Res.* 48:e64. doi: 10.1093/nar/gkaa294

Caro, F., Place, N. M., and Mekalanos, J. J. (2019). Analysis of lipoprotein transport depletion in *Vibrio cholerae* using CRISPRi. *Proc. Natl. Acad. Sci. U. S. A.* 116, 17013–17022. doi: 10.1073/pnas.1906158116

Cathomen, T., and Joung, J. K. (2008). Zinc-finger nucleases: the next generation emerges. *Mol. Ther.* 16, 1200–1207. doi: 10.1038/mt.2008.114

Cho, S. W., Kim, S., Kim, J. M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 230–232. doi: 10.1038/nbt.2507

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

Cress, B. F., Leitz, Q. D., Kim, D. C., Amore, T. D., Suzuki, J. Y., Linhardt, R. J., et al. (2017). CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production. *Microb. Cell Fact.* 16:10.

Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V., and Bikard, D. (2018). A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9. *Nat. Commun.* 9:1912.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607. doi: 10.1038/nature09886

Fontana, J., Dong, C., Ham, J. Y., Zalatan, J. G., and Carothers, J. M. (2018). Regulated expression of sgRNAs Tunes CRISPRi in *E. coli*. *Biotechnol. J.* 13:e1800069.

Guo, J., Wang, T., Guan, C., Liu, B., Luo, C., Xie, Z., et al. (2018). Improved sgRNA design in bacteria via genome-wide activity profiling. *Nucleic Acids Res.* 46, 7052–7069. doi: 10.1093/nar/gky572

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155

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

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

Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., and Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol.* 10, 841–851. doi: 10.4161/rna.24203

Kim, S. K., Seong, W., Han, G. H., Lee, D. H., and Lee, S. G. (2017). CRISPR interference-guided multiplex repression of endogenous competing pathway genes for redirecting metabolic flux in *Escherichia coli*. *Microb. Cell Fact.* 16:188.

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok i cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1156–1160. doi: 10.1073/pnas.93.3.1156

Lee, H. H., Ostrov, N., Wong, B. G., Gold, M. A., Khalil, A. S., and Church, G. M. (2019). Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi. *Nat. Microbiol.* 4, 1105–1113. doi: 10.1038/s41564-019-0423-8

Li, D., Lv, L., Chen, J. C., and Chen, G. Q. (2017). Controlling microbial PHB synthesis via CRISPRi. *Appl. Microbiol. Biotechnol.* 101, 5861–5867. doi: 10.1007/s00253-017-8374-6

Li, S., Jendresen, C. B., Landberg, J., Pedersen, L. E., Sonnenschein, N., Jensen, S. I., et al. (2020). Genome-wide CRISPRi-based identification of targets for decoupling growth from production. *ACS Synth. Biol.* 9, 1030–1040. doi: 10.1021/acssynbio.9b00143

Li, X. T., Jun, Y., Erickstad, M. J., Brown, S. D., Parks, A., Court, D. L., et al. (2016). tCRISPRi: tunable and reversible, one-step control of gene expression. *Sci. Rep.* 6:39076.

Liu, X., Gallay, C., Kjos, M., Domenech, A., Slager, J., van Kessel, S. P., et al. (2017). High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol. Syst. Biol.* 13:931. doi: 10.15252/msb.20167449

Liu, X., Kimmey, J. M., Matarazzo, L., de Bakker, V., Van Maele, L., Sirard, J. C., et al. (2021). Exploration of bacterial bottlenecks and *Streptococcus pneumoniae* pathogenesis by CRISPRi-Seq. *Cell Host Microbe* 29, 107–120. doi: 10.1016/j.chom.2020.10.001

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

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477.

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.

Man, S., Cheng, R., Miao, C., Gong, Q., Gu, Y., Lu, X., et al. (2011). Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res.* 39:e50. doi: 10.1093/nar/gkr034

Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., and Lee, S. Y. (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* 31, 170–174. doi: 10.1038/nbt.2461

Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., et al. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* 165, 1493–1506. doi: 10.1016/j.cell.2016.05.003

Peters, J. M., Koo, B. M., Patino, R., Heussler, G. E., Hearne, C. C., Qu, J., et al. (2019). Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. *Nat. Microbiol.* 4, 244–250. doi: 10.1038/s41564-018-0327-z

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

Qu, J., Prasad, N. K., Yu, M. A., Chen, S., Lyden, A., Herrera, N., et al. (2019). Modulating pathogenesis with Mobile-CRISPRi. *J. Bacteriol.* 201, e304–e319.

Rock, J. M., Hopkins, F. F., Chavez, A., Diallo, M., Chase, M. R., Gerrick, E. R., et al. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat. Microbiol.* 2:16274.

Rousset, F., Cui, L., Siouve, E., Becavin, C., Depardieu, F., and Bikard, D. (2018). Genome-wide CRISPR-dCas9 screens in *E. coli* identify essential genes and phage host factors. *PLoS Genet.* 14:e1007749. doi: 10.1371/journal.pgen.1007749

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR–a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6, 181–186. doi: 10.1038/nrmicro1793

van Opijnen, T., Bodi, K. L., and Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* 6, 767–772. doi: 10.1038/nmeth.1377

Wang, C., Cao, Y., Wang, Y., Sun, L., and Song, H. (2019). Enhancing surfactin production by using systematic CRISPRi repression to screen amino acid biosynthesis genes in *Bacillus subtilis*. *Microb Cell Fact.* 18:90.

Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., et al. (2018). Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.* 9:2475.

Wang, T., Wang, M., Zhang, Q., Cao, S., Li, X., Qi, Z., et al. (2019). Reversible gene expression control in Yersinia pestis by using an optimized CRISPR interference system. *Appl. Environ. Microbiol.* 85:e00097-9.

Wood, A. J., Lo, T. W., Zeitler, B., Pickle, C. S., Ralston, E. J., Lee, A. H., et al. (2011). Targeted genome editing across species using ZFNs and TALENs. *Science* 333:307.

Yang, G., Billings, G., Hubbard, T. P., Park, J. S., Yin Leung, K., Liu, Q., et al. (2017). Time-resolved transposon insertion sequencing reveals genome-wide fitness dynamics during infection. *mBio* 8:e015 81-17.

Yao, L., Shabestary, K., Bjork, S. M., Asplund-Samuelsson, J., Joensson, H. N., Jahn, M., et al. (2020). Pooled CRISPRi screening of the cyanobacterium Synechocystis sp PCC 6803 for enhanced industrial phenotypes. *Nat. Commun.* 11:1666.

Zhang, B., Liu, Z. Q., Liu, C., and Zheng, Y. G. (2016). Application of CRISPRi in *Corynebacterium glutamicum* for shikimic acid production. *Biotechnol. Lett.* 38, 2153–2161. doi: 10.1007/s10529-016-2207-z

Zhang, S., and Voigt, C. A. (2018). Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* 46, 11115–11125.

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.

frontiers in Microbiology

MINI REVIEW
published: 01 April 2021
doi: 10.3389/fmicb.2021.664299

Pruning and Tending Immune Memories: Spacer Dynamics in the CRISPR Array

*Sandra C. Garrett**

Department of Genetics and Genome Sciences, Institute for Systems Genomics, UConn Health, Farmington, CT, United States

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes) is a type of prokaryotic immune system that is unique in its ability to provide sequence-specific adaptive protection, which can be updated in response to new threats. CRISPR-Cas does this by storing fragments of DNA from invading genetic elements in an array interspersed with short repeats. The CRISPR array can be continuously updated through integration of new DNA fragments (termed spacers) at one end, but over time existing spacers become obsolete. To optimize immunity, spacer uptake, residency, and loss must be regulated. This mini-review summarizes what is known about how spacers are organized, maintained, and lost from CRISPR arrays.

Keywords: adaptation, spacer acquisition, repeat, array, CRISPR, spacer deletion

OPEN ACCESS

Edited by:
Muhammad Kamruzzaman, Westmead Institute for Medical Research, Australia

Reviewed by:
Malcolm F. White, University of St Andrews, United Kingdom
Alexander P. Hynes, McMaster University, Canada

***Correspondence:**
Sandra C. Garrett
garrett@uchc.edu;
sandracoralgarrett@gmail.com

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *04 February 2021*
Accepted: *12 March 2021*
Published: *01 April 2021*

Citation:
Garrett SC (2021) Pruning and Tending Immune Memories: Spacer Dynamics in the CRISPR Array. Front. Microbiol. 12:664299. doi: 10.3389/fmicb.2021.664299

INTRODUCTION

Prokaryotes have evolved a diverse repertoire of tools to restrict the proliferation of deleterious mobile genetic elements (Koonin et al., 2017). Uniquely among these tools, CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes) provides sequence-specific protection that can be updated in the face of novel threats, making it an adaptive immune system. CRISPR stores sequence information about potentially parasitic or harmful mobile genetic elements in an array (Barrangou et al., 2007) and uses that information to carry out targeted degradation of DNA or RNA, depending upon CRISPR type (Makarova et al., 2020). CRISPR-Cas systems are diverse and have been classified into two classes, six distinct types (I–VI), and at least 33 subtypes (Makarova et al., 2020), but certain characteristics are shared. All CRISPR arrays contain a series of direct repeats separated by short sequences called "spacers" which match DNA from previously encountered invaders (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). An upstream leader sequence regulates transcription of the array and mediates addition of new spacers (Jansen et al., 2002; Yosef et al., 2012; Wei et al., 2015; Alkhnbashi et al., 2016). In addition to the CRISPR array, there are usually nearby genes encoding CRISPR-associated (Cas) proteins, including nucleases.

After transcription, CRISPR array RNAs are processed into short guide RNAs (crRNAs) which associate with Cas nucleases to form a crRNA-guided effector complex (Hille et al., 2018). The crRNA base pairs with its complementary sequence in the target DNA or RNA (termed the "protospacer" since it corresponds to the invader nucleic acid that was originally captured and stored as a spacer) and leads to its degradation (interference). For DNA-targeting CRISPR systems, there must be a short activating sequence next to the target (called the Protospacer Adjacent Motif or PAM) for efficient interference (Deveau et al., 2008; Mojica et al., 2009; Shah et al., 2013). New

spacers are added to the array in a process called adaptation, wherein two proteins, Cas1 and Cas2, integrate fragments of DNA (McGinn and Marraffini, 2019) to produce new immune memories.

While the field has made great gains in understanding interference and adaptation in a wide range of organisms, many questions remain. For one, how are the individual immune memories in this heritable and adaptable system maintained over time? New spacers are continuously added in response to novel threats, but most arrays are less than 30 spacers long, suggesting that some immune memories are purged—which ones and how? This review will examine what we have learned about the dynamics of CRISPR arrays, with a focus on how immune memories (the spacers) are organized, maintained, and lost.

CRISPR Arrays Are Uniquely Organized Sequence Storage Banks

The most notable component of CRISPR-Cas systems is the repeat-spacer array, and the unusual structure of these elements was the first component of CRISPR-Cas to capture researchers attention as they studied nearby genes in *Escherichia coli* (Ishino et al., 1987; Nakata et al., 1989). Other types of repeats had been described in prokaryotic genomes, but in these new elements they found a novel layout: about a dozen direct repeats with loose dyad symmetry were arranged in a regularly spaced array (**Figure 1A**). The repeats were identical (or nearly identical) in sequence and length, while the intervening spacers had a common length but seemingly random sequence. The authors searched for and found the repeats in genomes of two other species of gram-negative bacteria and other groups found similar repeats in a range of bacteria and archaea (Groenen et al., 1993; Mojica et al., 1993; Mojica et al., 1995; Masepohl et al., 1996; Hoe et al., 1999). The broad distribution and surprisingly well-conserved layout suggested an important functional role for CRISPR arrays (Mojica et al., 2000; Jansen et al., 2002). That role was uncovered through a key observation about spacers: their sequences often matched DNA of mobile genetic elements like plasmids, phages, and prophages. Thus the CRISPR array appeared to be part of an immune system, with the spacer sequences acting as immune memories (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). This immune function was then confirmed directly. In a lab setting, cultures of a CRISPR-endowed strain of *Streptococcus thermophilus* were almost entirely killed off by lytic phage, but the small number of survivors (bacteriophage-insensitive mutants, BIM) all had at least one new spacer which matched the phage genome (Barrangou et al., 2007).

These and other studies showed that CRISPR-Cas could function as an immune system, and they also began to reveal general characteristics of how new spacers were acquired and stored. First, while studying bacteriophage-insensitive mutants it was noted that new spacers were added to one end of the array (Barrangou et al., 2007); this end contained the "leader," a 200–300 bp stretch of non-coding DNA (Jansen et al., 2002), which was later shown to regulate spacer uptake and array transcription (Pougach et al., 2010; Yosef et al., 2012). During uptake of a new spacer, the repeat was duplicated, so that an entire spacer-repeat unit was added (Barrangou et al., 2007). Later work showed that new spacer-repeat units could occasionally be added to the interior of the array, termed "ectopic" integration. In rare examples where ectopic integrations appear to outnumber leader-adjacent events, mutations in the leader were found and thought to cause the atypical localization (McGinn and Marraffini, 2016). Recently, ectopic integrations were reported in type II systems of *S. thermophilus* in the absence of leader mutations. While most (83%) integrations were leader-adjacent, the minority of ectopic events show that polarity is typical but not always absolute (Achigar et al., 2021).

Polarity of spacer uptake was not unique to lab–cultured organisms. Whenever CRISPR arrays from related strains were compared, a common pattern emerged: the greatest diversity of spacers was observed near the leader, with many of those spacers being unique to one strain or another, while the distal end of the array tended to have a series of spacers that was shared by many strains (**Figure 1A**; Pourcel et al., 2005; Lillestol et al., 2006; Horvath et al., 2008; Held et al., 2010; Lopez-Sanchez et al., 2012;

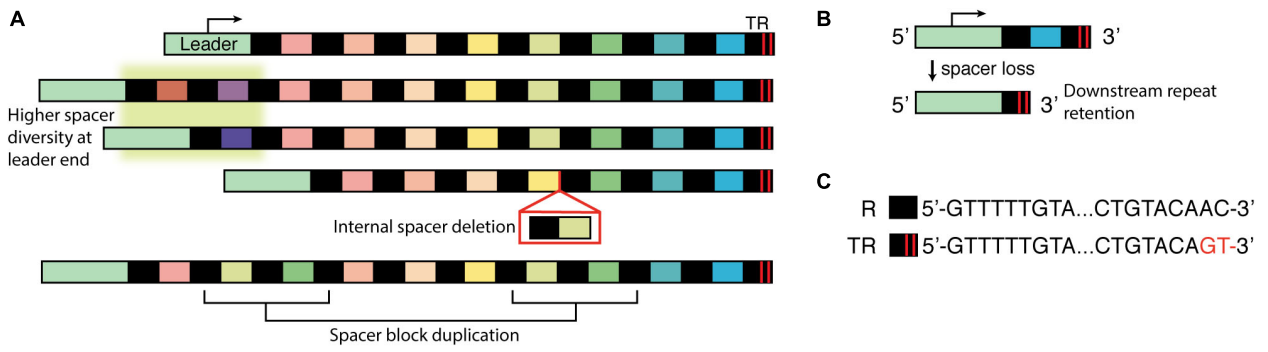


FIGURE 1 | Cartoon diagram of CRISPR array organization and key processes. **(A)** CRISPR arrays from related strains or isolates were compared; differences in spacer composition illustrated three key characteristics: highest spacer diversity is found at the leader-adjacent end of the array (highlighted in *green*), missing spacers in the array suggest spacer deletion events, and repeated blocks or repeated individual spacers suggests duplication events. **(B)** Sequencing showed that the downstream repeat is maintained after an upstream spacer-repeat unit deletion. **(C)** Example internal repeat (R) and terminal repeat (TR) for a type II-A CRISPR array from *Streptococcus thermophilus*. The 3′ end (leader distal) has two nucleotide substitutions, which disrupt dyad symmetry at the ends.

Lier et al., 2015; Rao et al., 2016). This pattern supported a polar and sequential process of spacer addition, with recent events near the leader and ancestral events at the distal end. A careful comparison of *Sulfolobus islandicus* isolates from a single hot spring lent particular support for this model: when two arrays shared non-identical spacers that likely arose from the same viral invader, the spacers were often in the same relative position within their respective arrays. The spacers' positions appeared to serve as a time stamp for the moment when the virus appeared in the spring and was captured into the CRISPR arrays (Held et al., 2010).

Second, it became clear that spacer-repeat units could be duplicated or deleted from the array. These changes were often observed in the middle of the array, while the distal end (termed "trailer" or "anchor" end) was typically conserved (Pourcel et al., 2005; Lillestol et al., 2006; Horvath et al., 2008; Lopez-Sanchez et al., 2012; Weinberger et al., 2012; Lam and Ye, 2019; Deecker and Ensminger, 2020). Evidence of losses or duplications was first inferred by comparing arrays from related strains; arrays that differed only by the absence of one or more contiguous spacers were thought to be the result of deletions (**Figure 1A**; Pourcel et al., 2005; Lillestol et al., 2006; Held et al., 2010; Gudbergsdottir et al., 2011; Lopez-Sanchez et al., 2012; Achigar et al., 2017). Spacer deletion was also sometimes detected while sequencing bacteriophage survivors (Deveau et al., 2008; Achigar et al., 2017). A minority of survivors both lost a contiguous block of existing spacers and added a new spacer against the experimental phage, leading some authors to suggest that "spacer deletion may occur concomitantly with the addition of new spacers" (Deveau et al., 2008). Repeated blocks of spacers were presumed to be duplications rather than independent adaptation events (Bolotin et al., 2005; Lillestol et al., 2006; Held et al., 2010; Lopez-Sanchez et al., 2012; Lier et al., 2015; Stout et al., 2018). While a second encounter with an old invader could conceivably lead to uptake of the same spacer twice, an identical series of spacers is unlikely.

Spacer deletion was particularly apparent in experiments wherein cultures were subjected to a selective pressure that favored failure of interference. For example, Jiang et al. (2013) introduced a conjugative plasmid encoding antibiotic resistance into *Staphylococcus epidermidis* RP62a, which had a type III-A CRISPR system and a spacer targeting the plasmid. When cultures were grown in the presence of antibiotics, interference against the plasmid resulted in 3–4 orders of magnitude fewer transconjugants as compared to controls. However, a few transconjugants were isolated and 13% of these had lost the plasmid-targeting spacer from their array. Other transconjugants had different mutations, all of which would disrupt CRISPR interference and thus allow the plasmid to persist and provide antibiotic resistance. Additional experiments suggested that these mutations arose spontaneously in the population rather than being induced by the selective pressure, and authors estimated that such mutations occurred in roughly one of 10^3 or 10^4 cells (Jiang et al., 2013).

In similar experiments with different organisms, spontaneous deletion of the targeting spacer was responsible for a larger share of escapees, in some cases occurring in more than 80% of the sequenced isolates (Gudbergsdottir et al., 2011; Lopez-Sanchez et al., 2012; Citorik et al., 2014; Rao et al., 2017; Stout et al., 2018; Canez et al., 2019). Deletion often included blocks of spacer-repeat units rather than only the targeting spacer (Gudbergsdottir et al., 2011; Lopez-Sanchez et al., 2012; Stout et al., 2018) and it was also sometimes associated with duplications of other non-targeting spacers (Lopez-Sanchez et al., 2012). As in the *S. epidermidis* work, rearrangements of the array were found even in the absence of selective pressure: for a strain of *Legionella pneumophila* bearing an engineered short array, roughly one of every 1,000–2,000 cells underwent a spontaneous spacer-repeat deletion (Rao et al., 2017). Sequencing revealed that the boundaries of the downstream repeat were maintained after the deletion, leading the authors to hypothesize that homologous recombination between repeats underlies array rearrangements (**Figure 1B**; Gudbergsdottir et al., 2011; Rao et al., 2017).

Mechanisms and Functions for Polarized Spacer Uptake

Polarized, leader-end addition of spacers was a reproducible observation and mechanisms soon emerged to show how it occurs. Cas1 and Cas2 are necessary for spacer uptake (Yosef et al., 2012) and are associated with all adaptation-active systems (Makarova et al., 2020). These two proteins are necessary and sufficient for *in vitro* integration (Nunez et al., 2015). In some organisms, Cas1 and Cas2 strongly favor integration at the leader-adjacent repeat and this bias is mediated by sequences in the leader (Wei et al., 2015; McGinn and Marraffini, 2016; Wright and Doudna, 2016; Xiao et al., 2017; Kim et al., 2019). But for other systems, *in vitro* experiments show that Cas1 and Cas2 alone will carry out integrations at other repeats in the array and even at repeat-like sequences outside the array (Nunez et al., 2015; Grainy et al., 2019). These same systems show polarized integration *in vivo* (Datsenko et al., 2012; Yosef et al., 2012; Shiimori et al., 2018), suggesting that additional factors can guide the reaction. In type I systems of *E. coli* and other bacteria, a protein called integration host factor (IHF) ensures polarization by binding to Cas1 and the leader (Nunez et al., 2016; Fagerlund et al., 2017; Wright et al., 2017). Other protein factors likely play a similar polarizing role in other organisms and await characterization (Rollie et al., 2018).

Several studies show that leader-adjacent integration is likely necessary for optimal immune function. A new spacer arises from contemporary mobile genetic elements, which likely represent the most current and therefore pressing threats for host cells. In addition, new spacers should be free of mismatches that accumulate for older spacers as their targets develop escape mutations (Deveau et al., 2008; Semenova et al., 2011; Cady et al., 2012; Rao et al., 2016). From that we could expect that leader-adjacent spacers would be prioritized for defense, and these spacers do indeed produce more robust interference (McGinn and Marraffini, 2016; Rao et al., 2016; Deecker and Ensminger, 2020). The mechanism underlying this difference is not entirely clear. Leader-adjacent spacers (and the crRNA's they encode) may be better expressed or more efficiently processed than downstream spacers. RNA sequencing data show more abundant crRNAs in the leader half of the array for many

CRISPR loci (Elmore et al., 2013; Carte et al., 2014; McGinn and Marraffini, 2016). Another possibility rests on the idea that individual crRNAs are essentially in competition to form a crRNP effector complex with less numerous Cas nucleases. As the first to be transcribed, leader-adjacent spacers may have a head start in a race to associate with Cas proteins and could suffer the least from the "dilution" effect of multiple spacers (Martynov et al., 2017). Experiments with a constructed mini-array lent support to the general idea of competition: a truncated array was created in a strain of *Legionella pneumoniae* by deleting all but the leader-adjacent spacer and its upstream and downstream repeats. This mini-array strain showed about 100-fold more plasmid targeting than the wildtype strain, which has 42 additional spacers downstream (Rao et al., 2017). Though the sequence and position was identical for the first spacer, loss of additional spacers dramatically increased its effectiveness.

CRISPR Arrays Vary in Length

In addition to influencing the polarity of spacer uptake, the dilution effect may also represent a functional constraint on the overall length of CRISPR arrays. Adaptation without spacer loss would presumably lead to ever-longer arrays, but among genomes sequenced so far, extremely long arrays are relatively rare. Arrays with greater than 100 spacers are observed; *Haliangium ochraceum* is a notable example, with a single array of 587 spacers and two other arrays measuring 189 and 36 spacers, respectively (Ivanova et al., 2010; Pourcel et al., 2020). However, a typical array contains fewer than 50 spacers in bacteria and fewer than 100 in archaea (Horvath et al., 2008; Mangericao et al., 2016; Pourcel et al., 2020). Array length does not appear to be limited by genome size (Pourcel et al., 2020) nor by cell resources: experimentally, lengthening an array by several spacers did not reduce fitness (Vale et al., 2015). Also, many genomes harbor more than one CRISPR system (up to 37 have been observed, in a species of *Actinoalloteichus*), and presumably the energy demands for a single 500 spacer array are similar to those for ten 50 spacer arrays. On the other hand, *cas* genes can have a fitness cost (Vale et al., 2015), so the observation that many organisms have evolved multiple short arrays suggests that array length is not limited by energetic costs of carrying extra spacers.

One hypothesis to explain array length patterns is that array size represents a tradeoff between the dilution effect described above and maintaining immunity to a range of potential threats, i.e., depth of immunity (Bradde et al., 2020). In turn, depth of immunity is balanced against the need to update the array frequently enough to contend with novel threats but not so frequently that the cell risks toxic auto-immunity (Stern et al., 2010; Vercoe et al., 2013; Weissman et al., 2018). Organisms may deal with dueling constraints by having multiple arrays, each with a different length and optimized depth of immunity (Weissman et al., 2018). This would imply that arrays can have different rates of both spacer uptake and loss. Regarding uptake, evidence already exists that adaptation efficiency varies among systems and can also change in response to certain cues like cell density (Hoyland-Kroghsbo et al., 2017; see Sternberg et al. (2016) for a general review of adaptation). Data on spacer loss is sparser, but at least one report suggests that the frequency of spacer loss can differ between systems in the same organism. Specifically, when otherwise identical plasmids with either a type I or type II mini-array were grown in *E. coli*, frequent spacer loss was observed for type I but not type II (Canez et al., 2019).

Spacer Turnover Is Not Strictly Chronological

As a spacer's residence time in an array increases, and it loses relevance, position, and sequence identity to its targets, we might expect selective pressures to no longer favor its maintenance. In that context, deletion events could be a useful means for shedding older spacers. However, multiple observations suggest that old spacers are not purged in a chronological manner and that mismatched or inefficient spacers may prove useful. A minority of older spacers can maintain identity to their protospacer targets, possibly due to stable or cyclical exposure to phages (Sun et al., 2016). We also now know that relatively ineffective spacers can participate in immunity through the process of primed adaptation. In short, priming occurs when a crRNP effector complex recognizes a protospacer target and then stimulates new spacer uptake using DNA located near that target (Datsenko et al., 2012; Swarts et al., 2012). Priming is observed even when interference is relatively inefficient, like when the protospacer does not have a canonical PAM or when there are mismatches between the crRNA and the protospacer, particularly in the "seed" region adjacent to the PAM (Semenova et al., 2011; Wiedenheft et al., 2011; Fineran et al., 2014; Li et al., 2014; Richter et al., 2014; Semenova et al., 2016; Garrett et al., 2020). Since primed adaptation tolerates these changes, a spacer that might otherwise be obsolete can contribute to CRISPR immunity by updating the CRISPR array. Experimentally, spacers in the middle of an array (*L. pneumophila*) were shown to give relatively inefficient interference but still effectively support priming (Deecker and Ensminger, 2020). Thus turnover of older spacers may not always appear steady or strictly chronological.

Many studies that demonstrated the polarity of spacer acquisition also described the relative stability of the array's trailer end (Lopez-Sanchez et al., 2012; Weinberger et al., 2012; Lam and Ye, 2019). Assuming sequential spacer uptake, we would expect these terminal spacers to be the oldest and thus the most likely to have lost protective potential. Phylogenetic relationships inferred from multilocus sequence typing supported the idea that terminal spacers are indeed ancestral (Lopez-Sanchez et al., 2012). These spacers should be lost if shedding is chronological, yet they are apparently deleted far less frequently than newer spacers toward the middle of the array. Therefore trailer end spacers may be maintained for reasons unrelated to their value in interference. The stability may be a simple outcome of fewer opportunities for recombination: an internal spacer can be lost through recombination involving any upstream or downstream repeat, but the last spacer would only be lost if recombination occurred at the terminal repeat. Terminal repeats may also be stabilized due to polymorphisms: in many systems the repeat sequences are identical throughout the array except at the end (Jansen et al., 2002; Bolotin et al., 2005; Pourcel et al., 2005; Horvath et al., 2008; Lopez-Sanchez et al., 2012;

Deecker and Ensminger, 2020; Refregier et al., 2020). Specifically, nucleotide substitutions or deletions are often found in the 3′ end of the terminal repeat (**Figure 1C**). Since identical repeats are most amenable to homologous recombination (Treangen et al., 2009), a trailer repeat without polymorphisms could potentially undergo recombination with the leader-adjacent repeat and eliminate the entire array, leaving only a copy of itself. Terminal repeat polymorphisms may thereby tend to reduce the likelihood of array collapse.

What would array collapse mean for immunity in the CRISPR locus? Experiments have confirmed that naïve adaptation can occur with a solitary leader-adjacent repeat (Yosef et al., 2012; Wei et al., 2015), which suggests that arrays could potentially be repopulated following a collapse, at least in laboratory conditions. Deecker and Ensminger (2020) also found evidence that "array collapse and repopulate" events occur naturally. First they showed that priming *in trans* could replenish a collapsed array in the lab: their strain of *L. pneumophila* naturally contains a type I-F system on both its chromosome and its endogenous plasmid. The chromosomal array was mutated to only contain the terminal repeat, mimicking a collapsed array. They transformed in a plasmid targeted by a spacer in the endogenous plasmid array and observed primed adaptation into the chromosomal collapsed array. The authors noted that patterns of repeat polymorphisms among naturally occurring *L. pneumophila* isolates looked like the replenished arrays they had created in the lab, suggesting this happens in nature. However, it remains unclear whether replenishment of a collapsed array is a universal phenomenon. If collapse is not well tolerated, terminal repeat polymorphisms may be functionally important in preventing it. On the other hand, if a system can readily bounce back from array collapse, terminal repeat polymorphisms may simply represent spontaneous mutations that persist because they are resistant to loss through recombination. Interestingly, many terminal repeat polymorphisms are nucleotide substitutions or truncations in the 3′ end, which partially disrupt the loose dyad symmetry of repeat ends (Jansen et al., 2002; Bolotin et al., 2005; Pourcel et al., 2005; Horvath et al., 2008; Lopez-Sanchez et al., 2012; Deecker and Ensminger, 2020). Since dyad symmetry is a frequently observed feature of repeats, one could speculate that loss of dyad symmetry helps stabilize terminal repeats.

From the studies discussed above, we can conclude that the trailer end of the array typically does not obey a pattern of chronological turnover. In an extreme example, spacer turnover across the entire array also bucks chronological turnover, even over the course of thousands of years. Savitskaya et al. (2017) acquired an intestinal microbiome sample from a well-preserved mammoth calf that was frozen for 42,000 years and they captured *E. coli* type I repeat-spacer amplicons by PCR. Reads primarily yielded data about individual spacers but a subset of reads were long enough to span two or three repeat-spacer units and these provided additional information about spacer order in the ancient arrays. Ancient spacers and spacer combinations were then compared to over 1,700 modern *E. coli* type I-E arrays from public databases. About 20% of the ancient spacers matched a modern spacer, and surprisingly those matches were positioned all over the modern arrays rather than being concentrated in the distal end. Trends for the two and three-spacer data were similar. This striking example demonstrated that for some systems, spacer order may not recapitulate a timeline of spacer acquisition.

Mechanisms for Spacer-Repeat Rearrangement

Repeats have often been associated with genome plasticity, and rearrangement of a repeat element like the CRISPR array is consistent with those observations. Repeats can undergo recombination through two general mechanisms: RecA-dependent homologous recombination and RecA-independent mechanisms like replication misalignment (slippage or slipped-strand mispairing) (Bzymek and Lovett, 2001; Treangen et al., 2009). In homologous recombination, RecA protein plays a key role as it binds and coats ssDNA and promotes strand exchange and annealing once it has found a region with sufficient sequence identity. The branched heteroduplex is then extended and resolved by, for example, RuvABC complex (Kowalczykowski, 2015). RecA-independent mechanisms also rely on homology but there are multiple distinguishing characteristics. First, RecA-independent mechanisms have shorter homology length requirements and are thought to be the primary source of recombination in prokaryotes for repeats that are less than about 200 bp (Bi and Liu, 1994; Lovett, 2004). Second, RecA-independent recombination is thought to involve the replication fork: in replication misalignment (the most well-described form of RepA-independent recombination), direct repeats mispair during replication, giving rise to duplications and deletions (Lovett, 2004). The frequency of this type of recombination increases with repeat length and identity and decreases as the spacing between repeats grows (Lovett, 2004). Disruption of replication can further promote misalignments and increase deletions (Michel, 2000).

Given that CRISPR repeats are short and closely spaced, a RecA-independent mechanism like replication misalignment could underlie array deletions and duplications. RecA is not necessary for adaptation (Ivancic-Bace et al., 2015; Radovcic et al., 2018). While there is not yet direct evidence for or against a role for RecA in CRISPR array rearrangements, a report about recombination in a CRISPR-derived system hints that it is not necessary. Ding et al. (2020) sought to improve the performance of dual guide RNA plasmids in CRISPR-based genome editing applications. The plasmids were designed to express two separate guide RNAs: each 20 bp guide spacer had an identical promoter (35 bp) upstream and identical "scaffold" (82 bp) downstream (the scaffold included 12 bp corresponding to the 5′ end of the CRISPR repeat, a 4 bp linker, and "tracr", the trans-activating crispr RNA important for forming a mature guide RNA). They found that the plasmid was extremely unstable: 73% had mutations, mostly deletions that excised either one of the two promoter-spacer-scaffold units. Changing the promoter to reduce the extent of homology did not eliminate deletions. The group also observed no reduction

in deletion frequency using strains with deleted or inactive RecA, showing that the process was not RecA dependent. Ultimately the group found that inverting one of the promoter-spacer-scaffold units was necessary to stop the deletion events. Interestingly, the authors found that growth and transformation conditions also influenced the frequency of deletions. Using electroporation instead of heat shock and culturing in rich growth medium both reduced (but not eliminated) deletions. They hypothesized that nutrient deprivation and DNA damage slow replication and thereby promote deletion through a replication mispairing mechanism on the lagging strand. These findings might not be directly applicable to native CRISPR arrays (for example, the homologous region in their engineered plasmid was longer—82 bp when different promoters were used), but they imply that natural spacer deletions could also be RecA-independent, possibly occurring through misalignment between repeats during replication (**Figure 2**).

DISCUSSION

Bringing together these different observations and experimental results, we can speculate on a general model for spacer dynamics in a CRISPR array: new spacers are added at the leader end of the array at some basic frequency, which varies among species, systems, and conditions. Rearrangement of the array is ongoing at some level, though the particular frequency is also variable among species and CRISRR-Cas classes and it may be modulated by as-of-yet unidentified factors and conditions. These rearrangements can lead to both deletions and duplications, and the interplay between spacer addition and loss determines array length and underpins the balance of immune depth, immune novelty, and crRNA dilution for that array. The terminal spacer-repeat unit rarely participates in rearrangements, potentially because of polymorphisms, so the array is maintained and the last spacer-repeat unit is stable. Together with adaptation events, rearrangements present immunogenic diversity on which selection can act. In most circumstances the dominant array form persists for generations, but the system is poised for change should conditions shift.

There is much more to learn about the dynamics and outcomes of spacer turnover. For one, it will be interesting to know how common array rearrangements are in different natural populations. Often these events were only detected because of a strong selective pressure against interference—how frequent are they in a native array under neutral conditions? Is the frequency consistent or does it vary with or independently from adaptation frequency? Since evidence suggests that the arrangements may not be equally common for all systems (Canez et al., 2019), it will be worthwhile to explore their frequency in multiple species and conditions. Long-read sequencing approaches may be particularly suitable for these experiments since they can capture the spacer composition of an entire array without the ambiguities inherent to assembled short reads.

Second, we have much to learn about the mechanism by which spacers are duplicated or deleted. The nature of CRISPR repeats and patterns of spacer loss are suggestive of rearrangement by recombination, but direct data are needed. For example, do deletions and duplications arise from RecA-independent mechanisms like misalignment in the replication fork? This would be supported by the results from the dual guide RNA plasmid experiments described above (Ding et al., 2020), and if experimentally confirmed, it could have interesting implications for immune diversity. If, for example, deletions primarily occur on the lagging strand during replication, we would expect them to be passed along to only one of two daughter cells. Since autoimmunity is thought to represent a fitness cost associated with CRISPR-Cas (Stern et al., 2010; Vercoe et al., 2013), replication fork deletion of new spacers may present a way to hedge against toxic self-targeting adaptation. If replication fork misalignment does underlie array rearrangements, are there factors or conditions that promote or inhibit the process and do they regulate spacer maintenance? And looking beyond the replication misalignment model, are there other enzymes or processes that can lead to spacer deletions or duplications? As we have learned more about adaptation and interference, points where the processes are modulated have been uncovered. Similarly, answering these and other questions about array dynamics may also help us uncover novel mechanisms that govern how existing spacers are managed to optimize immunity.

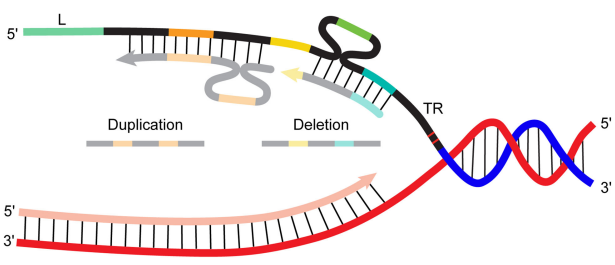


FIGURE 2 | Cartoon diagram of spacer-repeat duplication and deletion events created by misalignment of the nascent strand during replication. L, leader; TR, terminal repeat, all other repeats depicted in black, spacers depicted in color (*orange*, *yellow*, *green*, *aqua*).

AUTHOR CONTRIBUTIONS

SG devised and wrote the manuscript.

FUNDING

This work was supported by the National Institutes of Health (R35GM118140).

ACKNOWLEDGMENTS

I thank Brenton Graveley, members of the Graveley lab, and Michael Terns for helpful discussion during preparation of this work.

REFERENCES

Achigar, R., Magadan, A. H., Tremblay, D. M., Julia Pianzzola, M., and Moineau, S. (2017). Phage-host interactions in *Streptococcus thermophilus*: genome analysis of phages isolated in Uruguay and ectopic spacer acquisition in CRISPR array. *Sci. Rep.* 7:43438.

Achigar, R., Scarrone, M., Rousseau, G. M., Philippe, C., Machado, F., Duvos, V., et al. (2021). Ectopic spacer acquisition in *Streptococcus thermophilus* CRISPR3 array. *Microorganisms* 9:512. doi: 10.3390/microorganisms9030512

Alkhnbashi, O. S., Shah, S. A., Garrett, R. A., Saunders, S. J., Costa, F., and Backofen, R. (2016). Characterizing leader sequences of CRISPR loci. *Bioinformatics* 32, i576–i585.

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

Bi, X., and Liu, L. F. (1994). recA-independent and recA-dependent intramolecular plasmid recombination. Differential homology requirement and distance effect. *J. Mol. Biol.* 235, 414–423. doi: 10.1006/jmbi.1994.1002

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

Bradde, S., Nourmohammad, A., Goyal, S., and Balasubramanian, V. (2020). The size of the immune repertoire of bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 117, 5144–5151. doi: 10.1073/pnas.1903666117

Bzymek, M., and Lovett, S. T. (2001). Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8319–8325. doi: 10.1073/pnas.111008398

Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R., and O'Toole, G. A. (2012). The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* 194, 5728–5738. doi: 10.1128/jb.01184-12

Canez, C., Selle, K., Goh, Y. J., and Barrangou, R. (2019). Outcomes and characterization of chromosomal self-targeting by native CRISPR-Cas systems in *Streptococcus thermophilus*. *FEMS Microbiol. Lett.* 366:fnz105.

Carte, J., Christopher, R. T., Smith, J. T., Olson, S., Barrangou, R., Moineau, S., et al. (2014). The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.* 93, 98–112. doi: 10.1111/mmi.12644

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

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3:945.

Deecker, S. R., and Ensminger, A. W. (2020). Type I-F CRISPR-Cas distribution and array dynamics in *Legionella pneumophila*. *G3 (Bethesda)* 10, 1039–1050. doi: 10.1534/g3.119.400813

Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., et al. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400. doi: 10.1128/jb.01412-07

Ding, T., Huang, C., Liang, Z., Ma, X., Wang, N., and Huo, Y. X. (2020). Reversed paired-gRNA plasmid cloning strategy for efficient genome editing in *Escherichia coli*. *Microb. Cell Fact.* 19:63.

Elmore, J. R., Yokooji, Y., Sato, T., Olson, S., Glover, C. V. III, Graveley, B. R., et al. (2013). Programmable plasmid interference by the CRISPR-Cas system in *Thermococcus kodakarensis*. *RNA Biol.* 10, 828–840. doi: 10.4161/rna.24084

Fagerlund, R. D., Wilkinson, M. E., Klykov, O., Barendregt, A., Pearce, F. G., Kieper, S. N., et al. (2017). Spacer capture and integration by a type I-F Cas1-Cas2-3 CRISPR adaptation complex. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5122–E5128.

Fineran, P. C., Gerritzen, M. J., Suarez-Diez, M., Kunne, T., Boekhorst, J., van Hijum, S. A., et al. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E1629–E1638.

Garrett, S., Shiimori, M., Watts, E. A., Clark, L., Graveley, B. R., and Terns, M. P. (2020). Primed CRISPR DNA uptake in *Pyrococcus furiosus*. *Nucleic Acids Res.* 48, 6120–6135. doi: 10.1093/nar/gkaa381

Grainy, J., Garrett, S., Graveley, B. R., and PTerns, M. (2019). CRISPR repeat sequences and relative spacing specify DNA integration by *Pyrococcus furiosus* Cas1 and Cas2. *Nucleic Acids Res.* 47, 7518–7531. doi: 10.1093/nar/gkz548

Groenen, P. M., Bunschoten, A. E., van Soolingen, D., and van Embden, J. D. (1993). Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10, 1057–1065. doi: 10.1111/j.1365-2958.1993.tb00976.x

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

Held, N. L., Herrera, A., Cadillo-Quiroz, H., and Whitaker, R. J. (2010). CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLoS One* 5:e12988. doi: 10.1371/journal.pone.0012988

Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Hoe, N., Nakashima, K., Grigsby, D., Pan, X., Dou, S. J., Naidich, S., et al. (1999). Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. *Emerg. Infect. Dis.* 5, 254–263. doi: 10.3201/eid0502.990210

Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., et al. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1401–1412. doi: 10.1128/jb.01415-07

Hoyland-Kroghsbo, N. M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-Denomy, J., et al. (2017). Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proc. Natl. Acad. Sci. U.S.A.* 114, 131–135.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli,* and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi: 10.1128/jb.169.12.5429-5433.1987

Ivancic-Bace, I., Cass, S. D., Wearne, S. J., and Bolt, E. L. (2015). Different genome stability proteins underpin primed and naive adaptation in *E. coli* CRISPR-Cas immunity. *Nucleic Acids Res.* 43, 10821–10830. doi: 10.1093/nar/gkv1213

Ivanova, N., Daum, C., Lang, E., Abt, B., Kopitz, M., Saunders, E., et al. (2010). Complete genome sequence of *Haliangium ochraceum* type strain (SMP-2). *Stand. Genomic Sci.* 2, 96–106. doi: 10.4056/sigs.69.1277

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

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

Kim, J. G., Garrett, S., Wei, Y., Graveley, B. R., and Terns, M. P. (2019). CRISPR DNA elements controlling site-specific spacer integration and proper repeat length by a Type II CRISPR-Cas system. *Nucleic Acids Res.* 47, 8632–8648. doi: 10.1093/nar/gkz677

Koonin, E. V., Makarova, K. S., and Wolf, Y. I. (2017). Evolutionary genomics of defense systems in Archaea and Bacteria. *Annu. Rev. Microbiol.* 71, 233–261. doi: 10.1146/annurev-micro-090816-093830

Kowalczykowski, S. C. (2015). An overview of the molecular mechanisms of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* 7:a016410. doi: 10.1101/cshperspect.a016410

Lam, T. J., and Ye, Y. (2019). Long reads reveal the diversification and dynamics of CRISPR reservoir in microbiomes. *BMC Genomics* 20:567. doi: 10.1186/s12864-019-5922-8

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

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

Lillestol, R. K., Redder, P., Garrett, R. A., and Brugger, K. (2006). A putative viral defence mechanism in archaeal cells. *Archaea* 2, 59–72. doi: 10.1155/2006/542818

Lopez-Sanchez, M. J., Sauvage, E., Da Cunha, V., Clermont, D., Ratsima Hariniaina, E., Gonzalez-Zorn, B., et al. (2012). The highly dynamic CRISPR1

system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol. Microbiol.* 85, 1057–1071. doi: 10.1111/j.1365-2958.2012.08172.x

Lovett, S. T. (2004). Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. *Mol. Microbiol.* 52, 1243–1253. doi: 10.1111/j.1365-2958.2004.04076.x

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). 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

Mangericao, T. C., Peng, Z., and Zhang, X. (2016). Computational prediction of CRISPR cassettes in gut metagenome samples from Chinese type-2 diabetic patients and healthy controls. *BMC Syst. Biol.* 10(Suppl. 1):5. doi: 10.1186/s12918-015-0248-x

Martynov, A., Severinov, K., and Ispolatov, I. (2017). Optimal number of spacers in CRISPR arrays. *PLoS Comput. Biol.* 13:e1005891. doi: 10.1371/journal.pcbi.1005891

Masepohl, B., Gorlitz, K., and Bohme, H. (1996). Long tandemly repeated repetitive (LTRR) sequences in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Biochim. Biophys. Acta* 1307, 26–30. doi: 10.1016/0167-4781(96)00040-1

McGinn, J., and Marraffini, L. A. (2016). CRISPR-Cas systems optimize their immune response by specifying the site of spacer integration. *Mol. Cell* 64, 616–623. doi: 10.1016/j.molcel.2016.08.038

McGinn, J., and Marraffini, L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.* 17, 7–12. doi: 10.1038/s41579-018-0071-7

Michel, B. (2000). Replication fork arrest and DNA recombination. *Trends Biochem. Sci.* 25, 173–178. doi: 10.1016/s0968-0004(00)01560-7

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

Mojica, F. J., Diez-Villasenor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 36, 244–246. doi: 10.1046/j.1365-2958.2000.01838.x

Mojica, F. J., Ferrer, C., Juez, G., and Rodriguez-Valera, F. (1995). Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* 17, 85–93. doi: 10.1111/j.1365-2958.1995.mmi_17010085.x

Mojica, F. J., Juez, G., and Rodriguez-Valera, F. (1993). Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Mol. Microbiol.* 9, 613–621. doi: 10.1111/j.1365-2958.1993.tb01721.x

Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading)* 155, 733–740. doi: 10.1099/mic.0.023960-0

Nakata, A., Amemura, M., and Makino, K. (1989). Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *J. Bacteriol.* 171, 3553–3556. doi: 10.1128/jb.171.6.3553-3556.1989

Nunez, J. K., Bai, L., Harrington, L. B., Hinder, T. L., and Doudna, J. A. (2016). CRISPR immunological memory requires a host factor for specificity. *Mol. Cell* 62, 824–833. doi: 10.1016/j.molcel.2016.04.027

Nunez, J. K., Lee, A. S., Engelman, A., and Doudna, J. A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519, 193–198. doi: 10.1038/nature14237

Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L., et al. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.* 77, 1367–1379. doi: 10.1111/j.1365-2958.2010.07265.x

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

Pourcel, C., Touchon, M., Villeriot, N., Vernadet, J. P., Couvin, D., Toffano-Nioche, C., et al. (2020). CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res.* 48, D535–D544.

Radovcic, M., Killelea, T., Savitskaya, E., Wettstein, L., Bolt, E. L., and Ivancic-Bace, I. (2018). CRISPR-Cas adaptation in *Escherichia coli* requires RecBCD helicase but not nuclease activity, is independent of homologous recombination, and is antagonized by 5′ ssDNA exonucleases. *Nucleic Acids Res.* 46, 10173–10183.

Rao, C., Chin, D., and Ensminger, A. W. (2017). Priming in a permissive type I-C CRISPR-Cas system reveals distinct dynamics of spacer acquisition and loss. *RNA* 23, 1525–1538. doi: 10.1261/rna.062083.117

Rao, C., Guyard, C., Pelaz, C., Wasserscheid, J., Bondy-Denomy, J., Dewar, K., et al. (2016). Active and adaptive *Legionella* CRISPR-Cas reveals a recurrent challenge to the pathogen. *Cell. Microbiol.* 18, 1319–1338. doi: 10.1111/cmi.12586

Refregier, G., Sola, C., and Guyeux, C. (2020). Unexpected diversity of CRISPR unveils some evolutionary patterns of repeated sequences in *Mycobacterium tuberculosis*. *BMC Genomics* 21:841. doi: 10.1186/s12864-020-07178-6

Richter, C., Dy, R. L., McKenzie, R. E., Watson, B. N., Taylor, C., Chang, J. T., et al. (2014). Priming in the type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.* 42, 8516–8526. doi: 10.1093/nar/gku527

Rollie, C., Graham, S., Rouillon, C., and White, M. F. (2018). Prespacer processing and specific integration in a type I-A CRISPR system. *Nucleic Acids Res.* 46, 1007–1020. doi: 10.1093/nar/gkx1232

Savitskaya, E., Lopatina, A., Medvedeva, S., Kapustin, M., Shmakov, S., Tikhonov, A., et al. (2017). Dynamics of *Escherichia coli* type I-E CRISPR spacers over 42 000 years. *Mol. Ecol.* 26, 2019–2026. doi: 10.1111/mec.13961

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.

Semenova, E., Savitskaya, E., Musharova, O., Strotskaya, A., Vorontsova, D., Datsenko, K. A., et al. (2016). Highly efficient primed spacer acquisition from targets destroyed by the *Escherichia coli* type I-E CRISPR-Cas interfering complex. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7626–7631. doi: 10.1073/pnas.1602639113

Shah, S. A., Erdmann, S., Mojica, F. J., and Garrett, R. A. (2013). Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.* 10, 891–899. doi: 10.4161/rna.23764

Shiimori, M., Garrett, S. C., Graveley, B. R., and Terns, M. P. (2018). Cas4 nucleases define the PAM, length, and orientation of DNA fragments integrated at CRISPR loci. *Mol. Cell* 70, 814–824.e816.

Stern, A., Keren, L., Wurtzel, O., Amitai, G., and Sorek, R. (2010). Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26, 335–340. doi: 10.1016/j.tig.2010.05.008

Sternberg, S. H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas systems. *Mol. Cell* 61, 797–808.

Stout, E. A., Sanozky-Dawes, R., Goh, Y. J., Crawley, A. B., Klaenhammer, T. R., and Barrangou, R. (2018). Deletion-based escape of CRISPR-Cas9 targeting in *Lactobacillus gasseri*. *Microbiology (Reading)* 164, 1098–1111. doi: 10.1099/mic.0.000689

Sun, C. L., Thomas, B. C., Barrangou, R., and Banfield, J. F. (2016). Metagenomic reconstructions of bacterial CRISPR loci constrain population histories. *ISME J.* 10, 858–870. doi: 10.1038/ismej.2015.162

Swarts, D. C., Mosterd, C., van Passel, M. W., and Brouns, S. J. (2012). CRISPR interference directs strand specific spacer acquisition. *PLoS One* 7:e35888. doi: 10.1371/journal.pone.0035888

Treangen, T. J., Abraham, A. L., Touchon, M., and Rocha, E. P. (2009). Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiol. Rev.* 33, 539–571. doi: 10.1111/j.1574-6976.2009.00169.x

Vale, P. F., Lafforgue, G., Gatchitch, F., Gardan, R., Moineau, S., and Gandon, S. (2015). Costs of CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. *Proc. Biol. Sci.* 282:20151270. doi: 10.1098/rspb.2015.1270

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

Wei, Y., Chesne, M. T., Terns, R. M., and Terns, M. P. (2015). Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus*. *Nucleic Acids Res.* 43, 1749–1758. doi: 10.1093/nar/gku1407

Weinberger, A. D., Sun, C. L., Plucinski, M. M., Denef, V. J., Thomas, B. C., Horvath, P., et al. (2012). Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Comput. Biol.* 8:e1002475. doi: 10.1371/journal.pcbi.1002475

Weissman, J. L., Fagan, W. F., and Johnson, P. L. F. (2018). Selective maintenance of multiple CRISPR arrays across prokaryotes. *CRISPR J.* 1, 405–413. doi: 10.1089/crispr.2018.0034

Wiedenheft, B., van Duijn, E., Bultema, J. B., Waghmare, S. P., Zhou, K., Barendregt, A., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10092–10097. doi: 10.1073/pnas.1102716108

Wright, A. V., and Doudna, J. A. (2016). Protecting genome integrity during CRISPR immune adaptation. *Nat. Struct. Mol. Biol.* 23, 876–883. doi: 10.1038/nsmb.3289

Wright, A. V., Liu, J. J., Knott, G. J., Doxzen, K. W., Nogales, E., and Doudna, J. A. (2017). Structures of the CRISPR genome integration complex. *Science* 357, 1113–1118. doi: 10.1126/science.aao0679

Xiao, Y., Ng, S., Nam, K. H., and Ke, A. (2017). How type II CRISPR-Cas establish immunity through Cas1-Cas2-mediated spacer integration. *Nature* 550, 137–141. doi: 10.1038/nature24020

Yosef, I., Goren, M. G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* 40, 5569–5576. doi: 10.1093/nar/gks216

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

frontiers in Microbiology

PERSPECTIVE
published: 21 May 2021
doi: 10.3389/fmicb.2021.671522

Positioning Diverse Type IV Structures and Functions Within Class 1 CRISPR-Cas Systems

*Hannah N. Taylor[1], Eric Laderman[2], Matt Armbrust[1], Thomson Hallmark[1], Dylan Keiser[1], Joseph Bondy-Denomy[2] and Ryan N. Jackson[1]**

[1] Department of Chemistry and Biochemistry, Utah State University, Logan, UT, United States, [2] Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, United States

Type IV CRISPR systems encode CRISPR associated (Cas)-like proteins that combine with small RNAs to form multi-subunit ribonucleoprotein complexes. However, the lack of Cas nucleases, integrases, and other genetic features commonly observed in most CRISPR systems has made it difficult to predict type IV mechanisms of action and biological function. Here we summarize recent bioinformatic and experimental advancements that collectively provide the first glimpses into the function of specific type IV subtypes. We also provide a bioinformatic and structural analysis of type IV-specific proteins within the context of multi-subunit (class 1) CRISPR systems, informing future studies aimed at elucidating the function of these cryptic systems.

Keywords: CRISPR, Cas, type IV, Cas7, Cas6, DinG helicase, CysH

OPEN ACCESS

Edited by:
Aixin Yan,
The University of Hong Kong, Hong Kong, China

Reviewed by:
Lennart Randau,
University of Marburg, Germany
C. Martin Lawrence,
Montana State University, United States

****Correspondence:***
Ryan N. Jackson
ryan.jackson@usu.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *24 February 2021*
Accepted: *26 April 2021*
Published: *21 May 2021*

Citation:
Taylor HN, Laderman E, Armbrust M, Hallmark T, Keiser D, Bondy-Denomy J and Jackson RN (2021) Positioning Diverse Type IV Structures and Functions Within Class 1 CRISPR-Cas Systems. Front. Microbiol. 12:671522. doi: 10.3389/fmicb.2021.671522

INTRODUCTION

Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated (CRISPR-Cas) prokaryotic defense systems utilize Cas1 and Cas2 proteins, along with system-specific proteins such as Cas4, IHF, Csn2, and Cas9, to integrate foreign genetic material into the CRISPR locus, immunizing the cell against viruses and plasmids (Datsenko et al., 2012; Yosef et al., 2012; Nuñez et al., 2014, 2016; Heler et al., 2015; Rollie et al., 2015; Wang et al., 2015; Wei et al., 2015; Sternberg et al., 2016; Jackson S. A. et al., 2017; Kieper et al., 2018; Lee et al., 2018). To provide immunity, the CRISPR locus is transcribed and processed by RNA nucleases into CRISPR derived RNAs (crRNAs) (Brouns et al., 2008; Marraffini and Sontheimer, 2008; Haurwitz et al., 2010; Deltcheva et al., 2011). The crRNAs combine with Cas proteins to form ribonucleoprotein (RNP) complexes, which recognize and bind complementary nucleic acids. Binding induces cleavage of the foreign nucleic acid, protecting the cell (Brouns et al., 2008; Carte et al., 2008; Hale et al., 2008, 2009; Marraffini and Sontheimer, 2008; Garneau et al., 2010; Jackson R. N. et al., 2017; Hille et al., 2018).

Although all CRISPR systems use these general mechanisms to achieve immunity, the systems themselves are remarkably diverse, comprising two classes (1–2), six types (I–VI), and at least 33 subtypes (Yan et al., 2019; Makarova et al., 2020). In class 2 systems (types II, V, VI) a single Cas protein binds the crRNA to form the RNP complex, while class 1 RNP complexes (types I, III, IV) bind the crRNA with several proteins. Of the six CRISPR-Cas types, the least understood is type IV. Recent bioinformatic, biochemical, and structural studies of type IV CRISPR-Cas

systems have provided valuable insights into type IV system function. Here we compile known data on type IV systems, highlight recent advances in type IV system biology and biochemistry, and indicate questions concerning type IV systems that need to be addressed. Additionally, we provide phylogenetic analyses that suggest ancillary proteins associated with type IV systems have evolved Cas-specific functions.

TYPE IV SYSTEMS ARE MINIMAL, MOBILE CRISPR-CAS SYSTEMS

Distinguishable from other CRISPR-Cas systems, Type IV systems encode a distinct *cas7*-like gene (*csf2)*, lack adaptation genes, rarely encode an obvious nuclease, and are primarily found on plasmids (Koonin and Makarova, 2017, 2019; Pinilla-Redondo et al., 2019). These unique features of type IV systems have made it difficult to predict the function of type IV systems.

All type IV systems encode homologs of proteins known to form multi-subunit RNP complexes, explaining their class 1 designation. However, the presence of specific genes, gene arrangements, and differences in gene sequences have been used to further classify type IV systems into three distinct subtypes (IV-A, IV-B, and IV-C) (Makarova et al., 2011, 2015, 2020). Types IV-A, IV-B, and IV-C each contain a subtype-specific gene (*dinG, cysH-like,* and *cas10-like*, respectively) and subtype-specific features (**Figure 1A**). Type IV-A operons encode the three core type IV genes (*csf1, csf2,* and *csf3*), an endoribonuclease (*cas6*/*csf5*), a CRISPR array, and a putative helicase (*dinG*). Type IV-B operons encode the three core type IV genes and a *cas11-like* gene but lack a CRISPR locus. Additionally, type IV-B operons contain an ancillary gene, labeled *cysH-like* because the predicted secondary structure of its protein product resembles the CysH enzyme (Shmakov et al., 2018; Faure et al., 2019). Type IV-C systems encode *csf2* and *csf3*, but in place of the *csf1* gene they encode a *cas10-like* gene with a putative HD-nuclease domain. They also encode the *cas11-like* gene observed in IV-B systems, and sometimes a *cas6* RNA endonuclease and CRISPR array (Pinilla-Redondo et al., 2019; Makarova et al., 2020).

It is curious that the type IV systems that encode CRISPR loci do so in the absence of adaptation genes. It has been hypothesized that these type IV systems commandeer adaptation machinery from other CRISPR-Cas types to maintain their CRISPRs, similar to some type III systems (Staals et al., 2013, 2014; Elmore et al., 2015; Bernheim et al., 2020). Supporting this hypothesis, recent bioinformatic work showed that some type IV-A subtypes co-localize with certain type I systems (e.g., I-F, I-E), and that spacers found within co-localized type IV CRISPR loci appeared to be selected with the same criteria utilized by the type I system adaptation machinery [e.g., both I-E and IV-A protospacers are flanked with an 5′-AAG-3′ protospacer adjacent motif (PAM)] suggesting there may be functional cross-talk between these systems (Pinilla-Redondo et al., 2019). To confirm this proposed cooperation, *in vivo* and *in vitro* experimental work that examines adaptation in type IV systems with adaptation proteins from co-localized systems is needed.

THE *cas7*-LIKE GENE, *csf2*, DISTINGUISHES TYPE IV FROM OTHER CLASS 1 SYSTEMS

Initial bioinformatic analyses proposed *csf1* as the type IV *cas* signature gene (Makarova et al., 2015). However, some type IV systems lacking *csf1* have been identified, necessitating that the type IV *cas7* homolog, *csf2,* be used to classify type IV systems (Pinilla-Redondo et al., 2019). In type I and type III systems, Cas7-like proteins bind the crRNA guide within a helical backbone of a multi-subunit RNP complex and make direct interactions with other protein subunits (Jore et al., 2011; Lintner et al., 2011; Staals et al., 2013; Jackson et al., 2014; Mulepati et al., 2014; Osawa et al., 2015). Similarly, a recent cryo-EM structure of the type IV-B RNP complex revealed that Csf2 proteins bind RNA within a helical backbone, indicating a conserved function for Cas7-like proteins in all class 1 systems (Zhou et al., 2021). Despite this conservation, the sequence and structure of Csf2 is distinguishable from other Cas7 proteins (Makarova et al., 2011; **Supplementary Figure 1**). For example, when representative Cas7 sequences from all class 1 subtypes were aligned and a phylogenetic tree created, Csf2 sequences clustered on a separate branch from type I and type III Cas7 sequences (**Supplementary Figure 2A** and **Supplementary Methods**). Csf2 is distinct from other Cas7 homologs but appears to be most closely related to type III, supporting evolutionary hypotheses that type IV systems diverged from type III systems (Koonin and Makarova, 2019; Pinilla-Redondo et al., 2019; Makarova et al., 2020). Interestingly, an alignment of only Csf2 sequences shows clustering of Csf2 from each type IV subtype on its own branch, illustrating the intrinsic diversity of type IV subtypes and suggesting subtype-specific functional distinctions (**Supplementary Figure 2B**). It is worth noting that the type IV-B Csf2 subunit structure is most similar to the structure of the Cas7 homolog in type III-A systems, Csm3 (Zhou et al., 2021). Csm3 contains a catalytic aspartate that cleaves RNA targets (Tamulaitis et al., 2014). Alignment of target-bound Csm3 with Csf2 indicates that, although Csf2 also contains a conserved aspartate residue in a similar location, it is not in a position amenable for target cleavage (Zhou et al., 2021). Additional structural studies of type IV complexes bound to nucleic acid targets and complementary biochemical assays are needed to determine whether Csf2 is capable of RNA nuclease activity.

TYPE IV-A SYSTEMS ARE DEFENSE SYSTEMS WITH AN UNKNOWN MECHANISM OF ACTION INVOLVING A DinG HELICASE

Recently, a type IV-A system from *Pseudomonas aeruginosa* was shown to exhibit crRNA-guided defense against plasmids (Crowley et al., 2019), consistent with an analysis of type IV CRISPR spacers that suggested type IV-A systems disproportionately target plasmids (Pinilla-Redondo et al., 2019). Notably, earlier bioinformatic work indicated that

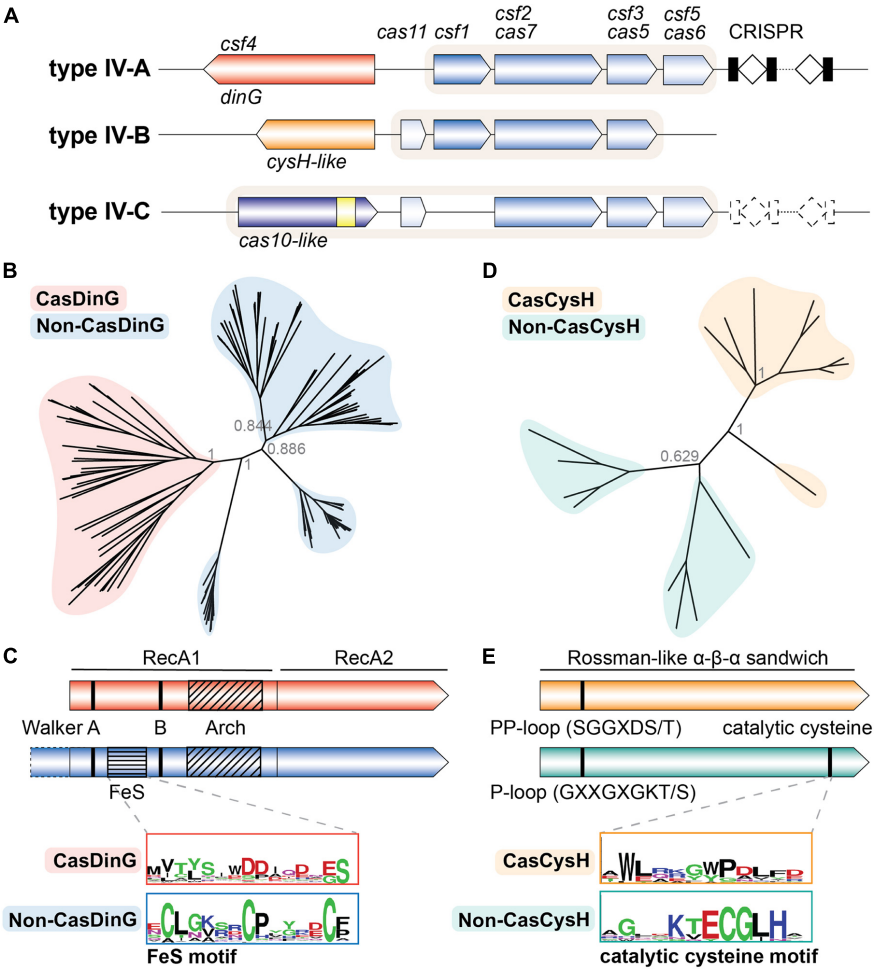


FIGURE 1 | The type IV Cas accessory proteins have evolved a Cas specific function. **(A)** Classification schematic of type IV CRISPR-Cas systems. A typical locus is represented for each type IV subtype. Dashed lines indicate components that are sometimes not encoded by the subtype. Shaded backgrounds highlight which gene products form the ribonucleoprotein (RNP) complex. The yellow square in the IV-C *cas10-like* large subunit represents an HD nuclease domain. **(B)** Phylogenetic tree of Cas- and non-CasDinG sequences. Posterior probabilities are shown. **(C)** Cartoons of Cas- and non-CasDinG sequences indicating positions of certain helicase motifs and domain architecture. Weblogos (Crooks, 2004) of the FeS cluster region in non-CasDinG (below, blue outline) and CasDinG (top, red outline) are shown. **(D)** Phylogenetic tree of Cas- and non-CasCysH sequences. Posterior probabilities are shown. **(E)** Cartoons of Cas- and non-CasCysH sequences. CasCysH is predicted to adopt the Rossman-like α - β - α fold observed in non-CasCysH structures. Positions and sequences of P- and PP-loops are indicated. Weblogos of the catalytic cysteine in non-CasCysH (bottom, teal outline) and CasCysH (top, orange outline) are shown.

many type IV-A spacers target viruses and prophage sequences encoding putative anti-CRISPRs, suggesting type IV-A systems also actively target viruses (Shmakov et al., 2017; Yin et al., 2019; Nobrega et al., 2020). However, direct data, such as viral plaque assays, are needed to confirm that type IV-A systems protect against viral attack.

Structural and biochemical work on a type IV-A complex from *Aromatoleum aromaticum* and IV-A Cas6 from *Mahella australiensis* demonstrated that the RNA endonuclease Csf5/Cas6 processes a crRNA upon which Csf1, Csf2, Csf3, and Csf5 form an RNP complex (Özcan et al., 2018; Taylor et al., 2019). At least three distinct crRNA processing endoribonucleases are encoded by Type IV-A systems (Cas6, Csf5, and Cas6e) (Makarova et al., 2020; **Supplementary Figure 3A**). Sequence alignments between biochemically characterized and putative type IV Csf5/Cas6 enzymes revealed Csf5 enzymes cleave RNA with arginine active site residues, while type IV Cas6 and Cas6e enzymes utilize histidine/tyrosine active site residues (**Supplementary Figures 3B, 4**). Despite these obvious differences in endoribonucleases, we hypothesize that in all type IV-A systems the Csf1, Csf2, Csf3, and Csf5/Cas6 proteins bind to the processed crRNA to form a multi-subunit complex that binds complementary nucleic acid.

It remains unclear whether type IV RNP complexes bind single stranded RNA [like the type III Csm and Cmr complexes (Hale et al., 2009; Samai et al., 2015)] or double stranded DNA [like the type I Cascade complexes (Brouns et al., 2008)] and how type IV complexes distinguish self from non-self targets. RNPs that target dsDNA usually rely on a protein-mediated binding event with a specific non-self sequence adjacent to

the complementary target, called a PAM (protospacer adjacent motif) (Mojica et al., 2009; Westra et al., 2012, 2013). PAM binding provides the energy for target duplex unwinding and interrogation of the DNA by the crRNA-guide. Work by Pinilla-Redondo et al. (2019) identified a consensus PAM (5′-AAG-3′) flanking protospacers targeted by a subset of type IV-A systems, suggesting type IV-A systems rely on PAM recognition to license binding. However, the consensus PAM may only reflect a preference of the acquisition machinery, which may explain why consensus PAM sequences have not been identified in all IV-A systems. Reliance on a specific PAM sequence for type IV-A RNP interference remains to be confirmed experimentally, but it should be noted that a promiscuous PAM recognition mechanism may indicate that the type IV complexes have evolved to accommodate the preferences of diverse Cas1 and Cas2 proteins that use different PAM sequences in spacer acquisition.

Interestingly, the structural similarities of the type IV-B complex to the type III Csm complex suggest that type IV complexes may target RNA (Zhou et al., 2021). Instead of recognizing a "non-self" PAM to license base pairing with a double-stranded DNA target, RNPs that bind RNA generally use a "self recognition" mechanism to distinguish self from non-self sequences (Marraffini and Sontheimer, 2010). Self-sequence located in the flanking regions of a bound RNA can base pair with the direct repeat of the crRNA disrupting downstream activation of effector nucleases (You et al., 2019). Self-sequences are inhibitory to overall immune function (Marraffini and Sontheimer, 2010; Elmore et al., 2016; Estrella et al., 2016; Kazlauskiene et al., 2016; Zhang et al., 2016; Han et al., 2017; Liu et al., 2017), but in some systems only a subset of non-self protospacer flanking sequences [called RNA-PAMs (rPAM) in type III systems or protospacer flanking sites (PFS) in type VI systems] are activating (Marraffini and Sontheimer, 2010; Abudayyeh et al., 2016; Elmore et al., 2016). We suspect that one or more Csf subunits may be responsible for PAM recognition to license DNA binding or rPAM recognition to activate immunity. We anticipate that *in vivo* PAM screens and biochemical binding assays with purified type IV-A RNPs will reveal the type IV-A self vs. non-self recognition mechanism.

Type IV-mediated plasmid clearance required all type IV-A system genes (*csf1, csf2, csf3, csf5,* and *dinG/csf4*) and a CRISPR containing a spacer complementary to a target plasmid sequence adjacent to a 5′-TTC-3′ PAM (Crowley et al., 2019). Because deleting the *dinG* gene or mutating the ATPase active site residues (DEAH-box) fully disrupted plasmid clearance, we hypothesize that RNP complex binding recruits the type IV-associated DinG (CasDinG) helicase to the bound target and CasDinG-mediated ATP binding and hydrolysis performs work, such as duplex unwinding, that is essential for plasmid clearance. Such a mechanism is similar to the more extensively studied type I Cas3 helicase-nuclease that unwinds and degrades dsDNA targets bound by the type I Cascade RNP complex (Beloglazova et al., 2011; Mulepati and Bailey, 2011; Sinkunas et al., 2011).

Both DinG and Cas3 classify as superfamily 2 helicases but, unlike Cas3, CasDinG proteins have no identifiable nuclease domain and have yet to be biochemically or structurally characterized (Fairman-Williams et al., 2010; Makarova et al., 2020). DinG helicases are generally involved in DNA recombination and repair, and are classified by amino acid sequence motifs involved in ATP binding and hydrolysis and nucleic acid binding and translocation (Lewis et al., 1992; Voloshin et al., 2003; Voloshin and Camerini-Otero, 2007; McRobbie et al., 2012; Wu and Brosh, 2012; Thakur et al., 2014; Cheng and Wigley, 2018). The motifs are located across two RecA helicase domains (**Supplementary Figure 5**). The first helicase domain also harbors two insertions, an iron sulfur cluster domain, and an arch domain, which are both important for duplex strand splitting (Ren et al., 2009; Peissert et al., 2020).

Since non-CasDinG helicases and their homologs have been extensively studied biochemically and structurally, we hypothesized that an in-depth comparison of CasDinG with non-CasDinG sequences would provide insight to CasDinG function. To investigate the relationship of CasDinG to other DinG helicases, we compiled CasDinG and non-CasDinG sequences from organisms containing a type IV-A system and generated a phylogenetic tree (**Supplementary Methods**; **Figure 1B**). Interestingly, CasDinG and non-CasDinG sequences clustered separately even when the sequences were retrieved from the same organism, suggesting CasDinG is functionally distinct from non-CasDinG. Notably, CasDinG helicases contain insertions within the first RecA domain of the same lengths as the iron-sulfur and arch insertions, but they lack homology with non-CasDinG sequences, including the residues important for coordinating the iron-sulfur cluster (**Figure 1C**). Sequence differences in these regions suggest these inserts may be a source of functional distinctions important for defense activities. Many functions for CasDinG have been hypothesized, including a role in displacing bound RNP complexes, cleaving bound targets with an unidentified nuclease activity (perhaps housed within an insert), or recruitment of endogenous nucleases to bound targets (Grodick et al., 2014). Notably, DinG helicases have been observed in a few type I and III systems (Dwarakanath et al., 2015; Makarova et al., 2020), indicating an evolutionary link and suggesting that some CasDinG activities essential for type IV immunity may have been co-opted by other class 1 systems.

In summary, recent bioinformatic and *in vivo* studies have indicated type IV-A systems protect prokaryotes from plasmids and viruses, but the mechanisms that underpin how the Csf RNP complex and CasDinG work together to provide immunity remain to be determined.

TYPE IV-B SYSTEMS ENCODE AN RNP COMPLEX OF UNKNOWN FUNCTION AND A SPECIALIZED CysH-LIKE PROTEIN WITH PUTATIVE ATP α-HYDROLASE ACTIVITY

Unlike type IV-A and IV-C subtypes, type IV-B systems lack a CRISPR locus and a crRNA processing enzyme, and are associated with an ancillary gene identified as *cysH-like* by the HHpred secondary structure prediction and alignment tool (Zimmermann et al., 2018; Makarova et al., 2020; **Figure 1A**).

A recent structural study recombinantly expressed and purified a *Mycobacterium sp. JS623* IV-B Csf RNP complex containing four type IV-B proteins (Csf1, Csf2, Csf3, and Cas11) (Zhou et al., 2021). Interestingly, RNA sequencing revealed the type IV-B Csf complex bound small heterogeneous RNAs, instead of co-expressed type I-E crRNAs from the *Mycobacterium sp. JS623* plasmid, suggesting a possible function other than CRISPR-mediated defense. A high resolution cryo-EM structure of the complex revealed several Csf2 subunits bind an RNA within a helical filament, while Cas11 subunits form a minor filament that contacts the larger filament at Csf2 dimer interfaces (Zhou et al., 2021). This structure of intertwined large and small protein filaments is similar to other class 1 RNP complexes, suggesting similar function as an RNA-guided complex that binds complementary targets (Jore et al., 2011; Lintner et al., 2011; Staals et al., 2013; Jackson et al., 2014; Mulepati et al., 2014; Osawa et al., 2015; **Supplementary Figure 6**).

Several observations are currently confounding an understanding of the type IV-B complex function. First, electron density for Csf1 and Csf3 subunits was not clearly observed within the structure, although SDS-PAGE indicated their presence in the purified complex. Thus, the structure and function of these important proteins remains unknown. Second, because the IV-B Csf complex bound heterogenous RNA, it remains unknown whether the Csf complex lacks sequence-specific preference for small RNAs or if the RNA(s) that the complex would normally bind were not available in the recombinant expression conditions. Finally, the role of the strictly conserved ancillary CysH-like protein and how it may interact with the complex is unknown.

The key to understanding the function of type IV-B systems likely lies with the uncharacterized, but ubiquitous, type IV-B accessory *cysH*-like gene (Shmakov et al., 2018; Faure et al., 2019). Typical CysH proteins are phosphoadenosine phosphosulfate (PAPS) reductases involved in sulfate assimilation. Structures reveal CysH proteins fit within a family of enzymes that adopt a Rossman-like α–β–α sandwich fold that binds nucleotides (InterPro IPR014729) (Blum et al., 2020). CysH proteins also contain a P-loop motif (GXXGXGKT/S consensus sequence) that binds nucleotide phosphates, and a conserved C-terminal cysteine that performs nucleophilic attack on the PAPS β-sulfate, hydrolyzing PAPS at the α-phosphate and forming a covalent thiosulfanate intermediate during sulfur reduction (Savage et al., 1997; Carroll et al., 2005; Chartron et al., 2007). Interestingly, the DndC protein from the recently discovered DND bacterial immune system also belongs to the PAPS reductase family, and uses a similar mechanism to incorporate sulfur into the backbone of chromosomal DNA through a disulfide cysteine (You et al., 2007; Wang et al., 2018; Faure et al., 2019). These phosphorothioate modifications serve as an epigenetic signature that allows the DND system to distinguish self from non-self DNA (Wang et al., 2018). The predicted structural homology between the type IV-B CysH (CasCysH) and DndC proteins justifies speculation that CasCysH proteins perform a similar function. However, a closer analysis of type IV-B CasCysH sequences suggests that if CasCysH does epigenetically modify DNA, it will not be through the formation of phosphorothioates. Although HHPred predicts CasCysH adopts a Rossman-like α–β–α sandwich fold, the catalytic cysteine important for sulfonate reduction in non-CasCysH and phosphothiolation of DNA by DndC is absent. Additionally, the P-loop sequence of CasCysH is more similar to the PP-motif (pyrophosphatase motif) (SGGXDS/T consensus sequence) observed in ATP PPases (Bork and Koonin, 1994; **Figure 1E**).

To better understand CasCysH activity and to explore the relationship between non-CasCysH and CasCysH proteins, sequences from organisms encoding both Cas- and non-CasCysH were aligned and phylogenetic trees determined. As was seen with Cas- and non-CasDinG, CasCysH sequences cluster separately from non-CasCysH sequences even when the sequences were retrieved from the same organisms (**Figure 1D**). Together, with our more in depth sequence analysis these differences suggest CasCysH evolved to preserve nucleotide binding without sulfonucleotide reduction.

Non-CasCysH enzymes fall within the larger classification of ATP α-hydrolases, which include N-type ATP PPases (Savage et al., 1997). Unlike non-CasCysH and DndC, N-type ATP PPases catalyze sequential reactions involving substrate AMPylation, instead of the formation of covalent enzyme substrate intermediates requiring nucleophilic attack from a catalytic cysteine (Chartron et al., 2007; Wang et al., 2018). The absence of a catalytic cysteine suggests that the role of CasCysH is to stabilize the AMPylation of specific substrates through catalysis of ATP α-hydrolase activity. We hypothesize that such an activity could be used to modify nucleic acids bound by the type IV-B RNP for immune system purposes, gene regulation, or the formation of secondary messengers. Future biochemical studies aimed at defining the function of CasCysH and its interactions with the IV-B Csf RNP complex will be critical for understanding type IV-B systems.

Several hypotheses exist concerning the function of type IV-B CRISPR-Cas systems. As they lack both a CRISPR array and an obvious nuclease it seems unlikely that type IV-B systems function as independent CRISPR-Cas defense systems (Makarova et al., 2011; Faure et al., 2019; Zhou et al., 2021). It has been suggested that type IV-B systems could bind crRNAs derived from other CRISPR systems, forming IV-B RNP complexes that perform RNA-guided defense (Makarova et al., 2011, 2015; Koonin and Makarova, 2019). As type IV systems are generally encoded on plasmids, such a crRNA scavenging system could be passed between organisms, acting as a mobile defense system. Interestingly, it was recently shown that sometimes type IV-B systems colocalize with specific class 1 systems, suggesting a cooperative function (Pinilla-Redondo et al., 2019). However, the same study showed that type IV-B systems are most often observed without other CRISPR systems, supporting an alternative hypothesis that proposes IV-B systems may protect plasmids from RNA-guided defense mechanisms by sponging up and inactivating small guide RNAs, including crRNAs (Koonin and Makarova, 2017; Faure et al., 2019; Pinilla-Redondo et al., 2019). Such an anti-guide-RNA activity could give plasmids containing

a type IV-B system a selective advantage (Shmakov et al., 2018; Koonin and Makarova, 2019). Although intriguing, neither of these hypotheses explain the role of the highly conserved ancillary protein CasCysH, suggesting the true function of IV-B systems may be more intricate than has so far been proposed.

THE NEWLY CLASSIFIED TYPE IV-C SYSTEM HIGHLIGHTS THE DIVERSE NATURE OF TYPE IV CRISPR-Cas SYSTEMS

Only recently did bioinformatics studies classify the subtype IV-C CRISPR-Cas system (Pinilla-Redondo et al., 2019; Makarova et al., 2020). Type IV-C systems lack a Csf1 subunit, and instead encode a Cas10-like subunit with an HD nuclease domain (**Figure 1A**). Type III CRISPR-Cas systems also encode Cas10, which is the large subunit of the RNP complex. In type III systems Cas10 has nuclease activity and synthesizes cyclic oligoadenylate second messengers (Jung et al., 2015; Kazlauskiene et al., 2017; Niewoehner et al., 2017). The type IV Cas10 contains an HD nuclease domain but not a nucleotide cyclase motif "GGDD," suggesting it has nuclease but not cyclic adenylate synthetase activity (Pinilla-Redondo et al., 2019). Interestingly, the HD domain motifs of type IV Cas10 are more similar to the HD domain of Cas3 than the type III Cas10 (Aravind and Koonin, 1998; Makarova et al., 2020). The presence of a *cas10-like* gene in type IV-C systems and the similarity between the type III-A and type IV-B RNP complexes support proposals that type IV and type III CRISPR-Cas systems share a common ancestor (Pinilla-Redondo et al., 2019; Makarova et al., 2020; Zhou et al., 2021). Experimental work is needed to better understand the function of these fascinating systems.

Several variants of type IV systems have been identified in bioinformatics studies and clinical samples which include

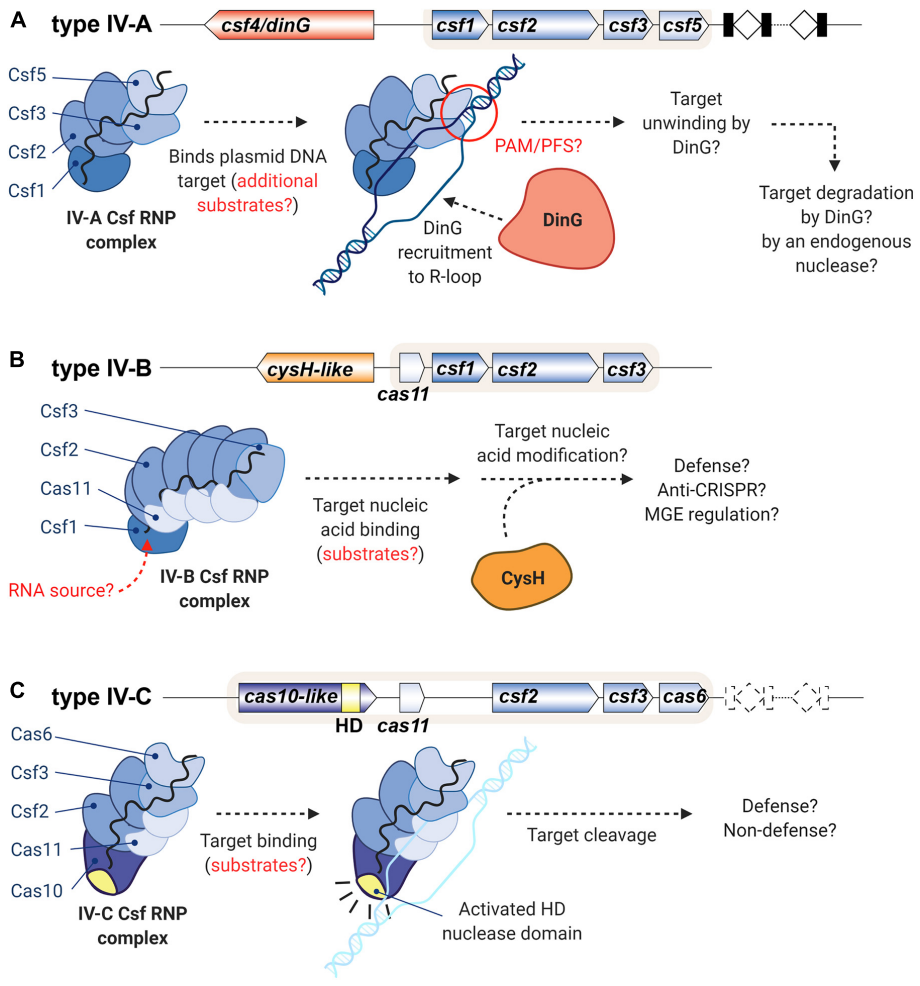


FIGURE 2 | Models of type IV system functions highlighting questions that remain to be answered. **(A)** IV-A RNP complexes likely bind DNA targets and recruit CasDinG for target unwinding and degradation. **(B)** IV-B RNP complexes likely interact with CasCysH to perform an unknown function. **(C)** The putative IV-C RNP complex likely binds a nucleic acid target and cleaves that target with the HD nuclease domain. Created with BioRender.com.

type IV systems; without a *csf1*, with a *csf1-csf3* fusion, with a *recD* helicase instead of *dinG*, and in association with IncH1b plasmids (Crowley et al., 2019; Pinilla-Redondo et al., 2019; Kamruzzaman and Iredell, 2020; Newire et al., 2020). We expect further study and analysis of these diverse systems will reveal unique mechanisms and functions.

UNANSWERED QUESTIONS CONCERNING TYPE IV BIOCHEMISTRY AND BIOLOGICAL FUNCTION

Throughout this perspective we have highlighted pressing questions concerning type IV CRISPR-Cas system structures and functions. Here we suggest models for the function of each type IV subtype and indicate areas which require further understanding. Type IV-A systems have been shown to form RNP complexes and prevent targeted plasmid transformation, but they have not been shown to target viruses nor is the mechanism of crRNA-guided defense clear (**Figure 2A**). Understanding the targets of the type IV-A system is critical to understanding the full scope of its defense activity. The presence of a helicase within the system suggests the need to unwind a duplex substrate. We suspect that the type IV-A system targets dsDNA, as it can defend against invasive plasmids (Crowley et al., 2019). However, CasDinG could also be important for unwinding duplex secondary structure within ssRNA targets or for targeting dsRNA phages (Poranen and Tuma, 2004). Remaining questions include the targeting parameters of the complex (DNA vs RNA, seed sequence, mismatch tolerance), how the complex distinguishes self from non-self, the role of CasDinG in immunity, and how targets are cleared without an identifiable nuclease domain within the system. We speculate that the IV-A RNP complex will bind to a dsDNA target and recruit CasDinG to the resulting R-loop, allowing CasDinG to unwind the target. To clear the target from the cell, we hypothesize that either an endogenous nuclease will degrade the unwound nucleic acid, or CasDinG harbors an intrinsic nuclease activity not predicted by the sequence.

The function of type IV-B Csf RNP complexes is still unknown (**Figure 2B**). Many questions of type IV-B system function will be answered as the source of the RNA component of the IV-B RNP complex is discovered and the function of the accessory protein CasCysH is understood. We propose that the Csf RNP complex will bind a nucleic acid target and recruit CasCysH to modify the nucleic acid via an ATP α-hydrolase activity.

No biochemical studies have been performed with type IV-C systems, to date. We hypothesize that the IV-C Csf proteins will form an RNP complex with a crRNA and the Cas10-like subunit (**Figure 2C**). The IV-C Csf RNP complex will bind a nucleic acid target complementary to the crRNA and the HD nuclease domain of the Cas10-like subunit will cleave the target. Some IV-C systems have a CRISPR and a crRNA processing endonuclease and others do not, suggesting some IV-C systems may serve a crRNA-guided defense function while others may employ Cas proteins to perform an entirely different, non-defense function. Future studies should seek to understand the role of Cas10 within the type IV Csf RNP complex and the overall function of type IV-C CRISPR-Cas systems.

To understand the function of type IV CRISPR-Cas systems, it is critical that we determine the structures and biochemical functions of the type IV subtype specific proteins: CasDinG, CasCysH, and Cas10-like. Phylogenetic trees suggest that the IV-A DinG and IV-B CysH have evolved to support a putative Cas specific function. The IV-C Cas10 also has a unique domain composition that likely supports a distinct function. We propose that, due to the different accessory proteins and subtype specific proteins encoded by the three subtypes, each type IV subtype will have a distinct mechanism of action and possibly distinct function. We highly anticipate future work detailing the mechanisms and functions of type IV RNP complexes and their accessory proteins.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

HT, TH, and RJ conceived and wrote the manuscript. HT, EL, MA, TH, and DK performed alignments and generated phylogenetic trees. All authors provided critical feedback on the manuscript.

FUNDING

JB-D and the Bondy-Denomy lab were supported by the UCSF Program for Breakthrough Biomedical Research funded in part by the Sandler Foundation, the Searle Fellowship, the Vallee Foundation, the Innovative Genomics Institute, and the NIH [DP5-OD021344 and R01GM127489]. Research in the Jackson Lab was supported by the National Institutes of Health National Institute of General Medical Sciences Maximizing Investigators' Research Award (R35GM138080).

ACKNOWLEDGMENTS

We thank members of the Jackson and Bondy-Denomy labs for helpful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.671522/full#supplementary-material

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:6299. doi: 10.1126/science.aaf5573

Aravind, L., and Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23, 469–472. doi: 10.1016/S0968-0004(98)01293-6

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

Bernheim, A., Bikard, D., Touchon, M., and Rocha, E. P. C. (2020). Atypical organizations and epistatic interactions of CRISPRs and cas clusters in genomes and their mobile genetic elements. *Nucl. Acids Res.* 48, 748–760. doi: 10.1093/nar/gkz1091

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., et al. (2020). The InterPro protein families and domains database: 20 years on. *Nucl. Acids Res.* 49, D344–D354. doi: 10.1093/nar/gkaa977

Bork, P., and Koonin, E. V. (1994). A P-loop-like motif in a widespread ATP pyrophosphatase domain: implications for the evolution of sequence motifs and enzyme activity. *Proteins* 20, 347–355. doi: 10.1002/prot.340200407

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., et al. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689

Carroll, K. S., Gao, H., Chen, H., Stout, C. D., Leary, J. A., and Bertozzi, C. R. (2005). A Conserved Mechanism for Sulfonucleotide Reduction. *PLoS Biol.* 3:e250. doi: 10.1371/journal.pbio.0030250

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

Chartron, J., Shiau, C., Stout, C. D., and Carroll, K. S. (2007). 3′-Phosphoadenosine-5′-phosphosulfate reductase in complex with thioredoxin: a structural snapshot in the catalytic cycle. *Biochemistry* 46, 3942–3951. doi: 10.1021/bi700130e

Cheng, K., and Wigley, D. B. (2018). DNA translocation mechanism of an XPD family helicase. *ELife* 7:e42400. doi: 10.7554/elife.42400.001

Crooks, G. E. (2004). WebLogo: a Sequence Logo Generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

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

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

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607. doi: 10.1038/nature09886

Dwarakanath, S., Brenzinger, S., Gleditzsch, D., Plagens, A., Klingl, A., Thormann, K., et al. (2015). Interference activity of a minimal Type I CRISPR–Cas system from Shewanella putrefaciens. *Nucl. Acids Res.* 43, 8913–8923. doi: 10.1093/nar/gkv882

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. *Nucl. Acids Res.* 43, 10353–10363. doi: 10.1093/nar/gkv1140

Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. *Genes Develop.* 30, 447–459. doi: 10.1101/gad.272153.115

Estrella, M. A., Kuo, F.-T., and Bailey, S. (2016). RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Develop.* 30, 460–470. doi: 10.1101/gad.273722.115

Fairman-Williams, M. E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr. Opinion Struct. Biol.* 20, 313–324. doi: 10.1016/j.sbi.2010.03.011

Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019). CRISPR–Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* 17, 513–525. doi: 10.1038/s41579-019-0204-7

Garneau, J. E., Dupuis, M. -È, 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

Grodick, M. A., Segal, H. M., Zwang, T. J., and Barton, J. K. (2014). DNA-Mediated Signaling by Proteins with 4Fe–4S Clusters Is Necessary for Genomic Integrity. *J. Am. Chem. Soc* 136, 6470–6478. doi: 10.1021/ja501973c

Hale, C., Kleppe, K., Terns, R. M., and Terns, M. P. (2008). Prokaryotic silencing (psi)RNAs in Pyrococcus furiosus. *RNA* 14, 2572–2579. doi: 10.1261/rna.1246808

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

Han, W., Li, Y., Deng, L., Feng, M., Peng, W., Hallstrøm, S., et al. (2017). A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. *Nucl. Acids Res.* 45, 1983–1993. doi: 10.1093/nar/gkw1274

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

Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., et al. (2015). Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature* 519, 199–202. doi: 10.1038/nature14245

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The Biology of CRISPR-Cas: backward and Forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Jackson, R. N., Golden, S. M., Erp, P. B. G., van, Carter, J., Westra, E. R., et al. (2014). Crystal structure of the CRISPR RNA–guided surveillance complex from *Escherichia coli*. *Science* 345, 1473–1479. doi: 10.1126/science.1256328

Jackson, R. N., van Erp, P. B., Sternberg, S. H., and Wiedenheft, B. (2017). Conformational regulation of CRISPR-associated nucleases. *Curr. Opinion Microbiol.* 37, 110–119. doi: 10.1016/j.mib.2017.05.010

Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., and Brouns, S. J. J. (2017). CRISPR-Cas: adapting to change. *Science* 356:eaal5056. doi: 10.1126/science.aal5056

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. Mole. Biol.* 18, 529–536. doi: 10.1038/nsmb.2019

Jung, T.-Y., An, Y., Park, K.-H., Lee, M.-H., Oh, B.-H., and Woo, E. (2015). Crystal Structure of the Csm1 Subunit of the Csm Complex and Its Single-Stranded DNA-Specific Nuclease Activity. *Structure* 23, 782–790. doi: 10.1016/j.str.2015.01.021

Kamruzzaman, M., and Iredell, J. R. (2020). CRISPR-Cas System in Antibiotic Resistance Plasmids in *Klebsiella pneumoniae*. *Front. Microbiol.* 10:2934. doi: 10.3389/fmicb.2019.02934

Kazlauskiene, M., Kostiuk, G., Venclovas, Č, Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* 357, 605–609. doi: 10.1126/science.aao0100

Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č, and Siksnys, V. (2016). Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: coupling DNA Degradation with the Target RNA Recognition. *Mol. Cell* 62, 295–306. doi: 10.1016/j.molcel.2016.03.024

Kieper, S. N., Almendros, C., Behler, J., McKenzie, R. E., Nobrega, F. L., Haagsma, A. C., et al. (2018). Cas4 Facilitates PAM-Compatible Spacer Selection during CRISPR Adaptation. *Cell Rep.* 22, 3377–3384. doi: 10.1016/j.celrep.2018.02.103

Koonin, E. V., and Makarova, K. S. (2017). Mobile Genetic Elements and Evolution of CRISPR-Cas Systems: all the Way There and Back. *Genome Biol. Evol.* 9, 2812–2825. doi: 10.1093/gbe/evx192

Koonin, E. V., and Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc B. Lond. Biol. Sci.* 374:20180087. doi: 10.1098/rstb.2018.0087

Lee, H., Zhou, Y., Taylor, D. W., and Sashital, D. G. (2018). Cas4-Dependent Prespacer Processing Ensures High-Fidelity Programming of CRISPR Arrays. *Mol. Cell* 70, 48–59.e5. doi: 10.1016/j.molcel.2018.03.003

Lewis, L. K., Jenkins, M. E., and Mount, D. W. (1992). Isolation of DNA damage-inducible promoters in *Escherichia coli*: regulation of polB (dinA), dinG, and dinH by LexA repressor. *J. Bacteriol.* 174, 3377–3385.

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. Y., Iavarone, A. T., and Doudna, J. A. (2017). RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. *PLoS One* 12:e0170552. doi: 10.1371/journal.pone.0170552

Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011). 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., Wolf, Y. I, Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). 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., 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

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

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

McRobbie, A.-M., Meyer, B., Rouillon, C., Petrovic-Stojanovska, B., Liu, H., and White, M. F. (2012). *Staphylococcus aureus* DinG, a helicase that has evolved into a nuclease. *Biochem. J.* 442, 77–84. doi: 10.1042/BJ20111903

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, 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

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., Héroux, 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

Newire, E., Aydin, A., Juma, S., Enne, V. I., and Roberts, A. P. (2020). Identification of a Type IV-A CRISPR-Cas System Located Exclusively on IncHI1B/IncFIB Plasmids in *Enterobacteriaceae*. *Front. Microbiol.* 11:1937. doi: 10.3389/fmicb.2020.01937

Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., Bigler, L., et al. (2017). Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548, 543–548. doi: 10.1038/nature23467

Nobrega, F. L., Walinga, H., Dutilh, B. E., and Brouns, S. J. J. (2020). Prophages are associated with extensive CRISPR–Cas auto-immunity. *Nucl. Acids Res.* 48, 12074–12084. doi: 10.1093/nar/gkaa1071

Nuñez, J. K., Bai, L., Harrington, L. B., Hinder, T. L., and Doudna, J. A. (2016). CRISPR Immunological Memory Requires a Host Factor for Specificity. *Mol. Cell* 62, 824–833. doi: 10.1016/j.molcel.2016.04.027

Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820

Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. *Mol. Cell* 58, 418–430. doi: 10.1016/j.molcel.2015.03.018

Özcan, A., Pausch, P., Linden, A., Wulf, A., Schühle, K., Heider, J., et al. (2018). Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat. Microbiol.* doi: 10.1038/s41564-018-0274-8

Peissert, S., Sauer, F., Grabarczyk, D. B., Braun, C., Sander, G., Poterszman, A., et al. (2020). In TFIIH the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat. Commun.* 11:1667. doi: 10.1038/s41467-020-15241-9

Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R. A., Randau, L., Sørensen, S. J., et al. (2019). Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. *Nucl. Acids Res.* 48, 2000–2012. doi: 10.1093/nar/gkz1197

Poranen, M. M., and Tuma, R. (2004). Self-assembly of double-stranded RNA bacteriophages. *Virus Res.* 101, 93–100. doi: 10.1016/j.virusres.2003.12.009

Ren, B., Duan, X., and Ding, H. (2009). Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster. *J. Biol. Chem.* 284, 4829–4835. doi: 10.1074/jbc.M807943200

Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L., and White, M. F. (2015). Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *ELife* 4:e08716.doi: 10.7554/eLife.08716

Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* 161, 1164–1174. doi: 10.1016/j.cell.2015.04.027

Savage, H., Montoya, G., Svensson, C., Schwenn, J. D., and Sinning, I. (1997). Crystal structure of phosphoadenylyl sulphate (PAPS) reductase: a new family of adenine nucleotide α hydrolases. *Structure* 5, 895–906. doi: 10.1016/S0969-2126(97)00244-X

Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Severinov, K. V., and Koonin, E. V. (2018). Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl Acad. Sci.* 115, E5307–E5316. doi: 10.1073/pnas.1803440115

Shmakov, S. A., Sitnik, V., Makarova, K. S., Wolf, Y. I., Severinov, K. V., and Koonin, E. V. (2017). The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *mBio* 8, 01397-17. doi: 10.1128/mBio.01397-17

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: cas3 nuclease/helicase. *EMBO J.* 30, 1335–1342. doi: 10.1038/emboj.2011.41

Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., et al. (2013). Structure and Activity of the RNA-Targeting Type III-B CRISPR-Cas Complex of Thermus thermophilus. *Mol. Cell* 52, 135–145. doi: 10.1016/j.molcel.2013.09.013

Staals, R. H. J., Zhu, Y., Taylor, D. W., Kornfeld, J. E., Sharma, K., Barendregt, A., et al. (2014). RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. *Mol. Cell* 56, 518–530. doi: 10.1016/j.molcel.2014.10.005

Sternberg, S. H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Mol. Cell* 61, 797–808. doi: 10.1016/j.molcel.2016.01.030

Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, È, Nwokeoji, A. O., Dickman, M. J., et al. (2014). Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. *Mol. Cell* 56, 506–517. doi: 10.1016/j.molcel.2014.09.027

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

Thakur, R. S., Desingu, A., Basavaraju, S., Subramanya, S., Rao, D. N., and Nagaraju, G. (2014). Mycobacterium tuberculosis DinG Is a Structure-specific Helicase That Unwinds G4 DNA Implications for Targeting G4 DNA as a Novel Therapeutic Approach. *J. Biol. Chem.* 289, 25112–25136. doi: 10.1074/jbc.M114.563569

Voloshin, O. N., and Camerini-Otero, R. D. (2007). The DinG Protein from *Escherichia coli* Is a Structure-specific Helicase. *J. Biol. Chem.* 282, 18437–18447. doi: 10.1074/jbc.M700376200

Voloshin, O. N., Vanevski, F., Khil, P. P., and Camerini-Otero, R. D. (2003). Characterization of the DNA Damage-inducible Helicase DinG from *Escherichia coli*. *J. Biol. Chem.* 278, 28284–28293. doi: 10.1074/jbc.M301188200

Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., et al. (2015). Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell* 163, 840–853. doi: 10.1016/j.cell.2015.10.008

Wang, L., Jiang, S., Deng, Z., Dedon, P. C., and Chen, S. (2018). DNA phosphorothioate modification—A new multi-functional epigenetic system in bacteria. *FEMS Microbiol. Rev.* 43, 109–122. doi: 10.1093/femsre/fuy036

Wei, Y., Terns, R. M., and Terns, M. P. (2015). Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. *Genes Develop.* 29, 356–361. doi: 10.1101/gad.257550.114

Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K., et al. (2013). Type I-E CRISPR-Cas Systems Discriminate Target from Non-Target DNA through Base Pairing-Independent PAM Recognition. *PLoS Gene.* 9:e1003742. doi: 10.1371/journal.pgen.1003742

Westra, E. R., van Erp, P. B. G., Künne, T., Wong, S. P., Staals, R. H. J., Seegers, C. L. C., 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

Wu, Y., and Brosh, R. M. Jr. (2012). DNA helicase and helicase–nuclease enzymes with a conserved iron–sulfur cluster. *Nucl. Acids Res.* 40, 4247–4260. doi: 10.1093/nar/gks039

Yan, W. X., Hunnewell, P., Alfonse, L. E., Carte, J. M., Keston-Smith, E., Sothiselvam, S., et al. (2019). Functionally diverse type V CRISPR-Cas systems. *Science* 363, 88–91. doi: 10.1126/science.aav7271

Yin, Y., Yang, B., and Entwistle, S. (2019). Bioinformatics Identification of Anti-CRISPR Loci by Using Homology, Guilt-by-Association, and CRISPR Self-Targeting Spacer Approaches. *mSystems* 4, e455–e419. doi: 10.1128/mSystems.00455-19

Yosef, I., Goren, M. G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucl. Acids Res.* 40, 5569–5576. doi: 10.1093/nar/gks216

You, D., Wang, L., Yao, F., Zhou, X., and Deng, Z. (2007). A Novel DNA Modification by Sulfur: dndA Is a NifS-like Cysteine Desulfurase Capable of Assembling DndC as an Iron-Sulfur Cluster Protein in *Streptomyces lividans* †. *Biochemistry* 46, 6126–6133. doi: 10.1021/bi602615k

You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., et al. (2019). Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. *Cell* 176, 239–253. doi: 10.1016/j.cell.2018.10.052

Zhang, J., Graham, S., Tello, A., Liu, H., and White, M. F. (2016). Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. *Nucl. Acids Res.* 44, 1789–1799. doi: 10.1093/nar/gkw020

Zhou, Y., Bravo, J. P. K., Taylor, H. N., Steens, J. A., Jackson, R. N., Staals, R. H. J., et al. (2021). Structure of a type IV CRISPR-Cas ribonucleoprotein complex. *iScience* 24:102201. doi: 10.1016/j.isci.2021.102201

Zimmermann, L., Stephens, S., Seung-Zin, N., Rau, D., Kubler, J., Lozajic, M., et al. (2018). A Completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243. doi10.1016/j.jmb.2017.12.007

Conflict of Interest: JB-D is a scientific advisory board member of SNIPR Biome and Excision Biotherapeutics, and a scientific advisory board member and co-founder of Acrigen Biosciences.

The remaining 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.

frontiers in Microbiology

MINI REVIEW
published: 28 May 2021
doi: 10.3389/fmicb.2021.657981

The Challenge of CRISPR-Cas Toward Bioethics

*Luis Uriel Gonzalez-Avila[1], Juan Manuel Vega-López[1], Leda Ivonne Pelcastre-Rodríguez[1], Omar Alejandro Cabrero-Martínez[1], Cecilia Hernández-Cortez[2] and Graciela Castro-Escarpulli[1]**

[1] Laboratorio de Investigación Clínica y Ambiental, Departamento de Microbiología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de Mexico, Mexico, [2] Laboratorio de Bioquímica Microbiana, Departamento de Microbiología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de Mexico, Mexico

Since determining the structure of the DNA double helix, the study of genes and genomes has revolutionized contemporary science; with the decoding of the human genome, new findings have been achieved, including the ability that humans have developed to modify genetic sequences *in vitro*. The discovery of gene modification mechanisms, such as the CRISPR-Cas system (Clustered Regularly Interspaced Short Palindromic Repeats) and Cas (CRISPR associated). Derived from the latest discoveries in genetics, the idea that science has no limits has exploded. However, improvements in genetic engineering allowed access to new possibilities to save lives or generate new treatment options for diseases that are not treatable by using genes and their modification in the genome. With this greater knowledge, the immediate question is who governs the limits of genetic science? The first answer would be the intervention of a legislative branch, with adequate scientific advice, from which the logical answer, bioethics, should result. This term was introduced for the first time by Van Rensselaer Potter, who in 1970 combined the Greek words bios and ethos, *Bio-Ethik*, which determined the study of the morality of human behavior in science. The approach to this term was introduced to avoid the natural tension that results from the scientific technical development and the ethics of limits. Therefore, associating the use of biotechnology through the CRISPR-Cas system and the regulation through bioethics, aims to monitor the use of techniques and technology, with benefits for humanity, without altering fundamental rights, acting with moral and ethical principles.

Keywords: CRISPR-Cas, bioethics, laws, genetic, biotechnology

OPEN ACCESS

Edited by:
Haike Antelmann,
Freie Universität Berlin, Germany

Reviewed by:
Gert Bange,
University of Marburg, Germany
Yoshizumi Ishino,
Kyushu University, Japan
Runhua Han,
University of Texas at Austin, United States

****Correspondence:***
Graciela Castro-Escarpulli
chelacastro@hotmail.com

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *24 January 2021*
Accepted: *03 May 2021*
Published: *28 May 2021*

Citation:
Gonzalez-Avila LU, Vega-López JM, Pelcastre-Rodríguez LI, Cabrero-Martínez OA, Hernández-Cortez C and Castro-Escarpulli G (2021) The Challenge of CRISPR-Cas Toward Bioethics. Front. Microbiol. 12:657981. doi: 10.3389/fmicb.2021.657981

INTRODUCTION

Since the discovery of the DNA structure described by Watson and Crick in 1953, the generation of knowledge about the molecular genetic bases began. It was determined that the double helix contained all the genetic information of the individuals made up of the four bases, adenine, thymine, guanine, and cytosine. After this event, in the field of genetics, the human genome was sequenced, discovering that it is made up of 3 billion base pairs, which oversees the production of millions of different proteins with the help of the complex cellular system, which allows the body to function. These advances have allowed the development of gene therapy, through which it is sought to interfere in gene expression through corrective manipulation based on sequence cutting and pasting techniques (Pérez, 2006; Yabar, 2019; Espinosa and Hernández-Hernández, 2020).

Advances in genetic engineering have been advancing, and proposals for innovations and simplification of techniques, as more details about DNA are known, allow the study and understanding of the complex genomic system of expression and the shutdown of the genes. Molecular techniques have sought to correct the damage in the sequence of the carrier of a disease encoded in the genome; however, they have not achieved their objective, since there is no absolute control over the damage that can occur in the carrier, trying to prevent it from the damage inherited to the offspring. The use of these techniques involves diverse and complex techniques *in vivo* and *in vitro*, the mechanisms used are mainly based on the use of vectors that seek to introduce a specific or modified gene, which is capable of being transcribed and the mRNA is produced to be translated (You et al., 2019; Jamal et al., 2020).

The three principles of bioethics initially proposed in the Belmont report in 1978, were beneficence, autonomy of patients, and justice. Later, in the work of Beauchamp and Childres Principles of Biomedical Ethics, they added the fourth principle, which they called non-maleficence. These principles that obey a marinist philosophical reflection, which was initially promoted by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, constituted the study of ethical issues related to biomedical research (Gómez-Tatay and Aznar, 2019; Shkomova, 2019; Schupmann et al., 2020).

In this concept, non-maleficence, would highlight the premise of *Primum non nocere*, translated as "first, do no harm," based upon the studies on the corresponding criteria. Avoiding the improper use of the sequences, especially the molecular biology techniques that are used for this. Those that use restriction enzymes, cloning of sequences in plasmids, integrons, the use of nucleases or the recently described CRISPR Cas, which, since its discovery, it generated interest in the scientific community due to the rarity and complexity of the system, contemporary medicine, and even technology are not allowed (Capella, 2016; Cribbs and Perera, 2017; Noll, 2019).

CRISPR-Cas OVERVIEW

The acronym CRISPR comes from Clustered Regularly Interspaced Short Palindromic Repeats, and the second part Cas refers to nuclease-like proteins that are associated with the CRISPR system (CRISPR associated system) (Capella, 2016; Hille et al., 2018).

The first CRISPR-Cas systems being detected over 30 years ago in *Escherichia coli* (Ishino et al., 1987) and with the acronym of the system being defined on the early 2000s (Jansen et al., 2002); the overall study of these systems has become widely popular due to their properties and multiple applications. The CRISPR-Cas genomic loci consist of a CRISPR array composed of direct repeats with unique spacers between them and the Cas genes, the number of these arrays that one genome can harbor ranges from 1 to 18, while the number of repeat units in one array ranges from 2 to 374 (Marraffini and Sontheimer, 2010).

These systems, more widely known as genome engineering tools, achieve immunity by incorporating fragments of foreign nucleic acids into the CRISPR arrays, enabling a series of proteins to sense by base-pair complementarity to perform the cleavage of the specific DNA or RNA sequences from the exogenic elements (Makarova et al., 2020).

The immune response creates and keeps updating a molecular file of encounters with foreign nucleic acids in the form of spacers; sequences of typically 32–38 nucleotides (nt) of length, ranging from 21 to 72 nt (Barrangou and Marraffini, 2014). These spacers are subsequently used to protect the bacterium or archaeon against new infections with a similar agent (Faure et al., 2019). Acquired spacers in the adaptation stage are then transcribed and processed into crRNAs (CRISPR RNAs) in the expression stage to allow the start of the final interference stage, in which the crRNA is used to recognize the complementary or partially complementary sequence of the spacer present in the invading mobile genetic element (MGE), this is followed by the cleavage and inactivation of said element by either one or more Cas nucleases (Faure et al., 2019).

The defensive strategy consists of the generation of the guide RNA that is an exact copy of the viral DNA, said RNA sequence will function as a guide for the Cas protein for the identification of the genetic material of the virus. By means of the complementary sequences of DNA and gRNA, they hybridize, concluding with the cut made by the Cas protein for viral inactivation (Mojica et al., 2005).

In general, the operation can be understood in three phases; the first consists of adaptation, in which foreign genetic material is incorporated into the locus spacers to save said sequences for future attacks. The incorporation of these sequences is carried out by horizontal gene transfer, degrading the foreign DNA, leaving it as new spacers. The second phase consists of the transcription of the CRISPR Cas, generating a precursor (CRISPR-RNA or pre-crRNA), which is processed and generates the crRNAs that are complementary to the foreign DNA sequence. In the third phase, commonly known as interference, the Cas proteins, using guide crRNAs, detect foreign sequences and degrade them (Makarova et al., 2015; Capella, 2016; Mojica and Rodriguez-Valera, 2016).

The CRISPR revolution has been made possible by the identification of the right enzymatic systems that simplify methodologies to exploit the potential of CRISPR-Cas systems, in a similar fashion to the development of the polymerase chain reaction (PCR) (Ishino et al., 2018). Due to the complexity and potential biotechnological application, research has increased exponentially, allowing studies on the genetic manipulation of species, modification of cell lines, and the creation of new mutants (**Figure 1**). One of the most important is the genetic modification of patients affected by a disease, but who sets the limits for these scientific advances? (Caballero-Hernández et al., 2017; Hirsch et al., 2019).

CRISPR-Cas Technology Advancements

Science from its conception tries to generate knowledge and discoveries that allow a benefit for humanity. The possibility of having a cure for almost any genetically based disease is undeniable, however, as in all stories, there is a good and a bad

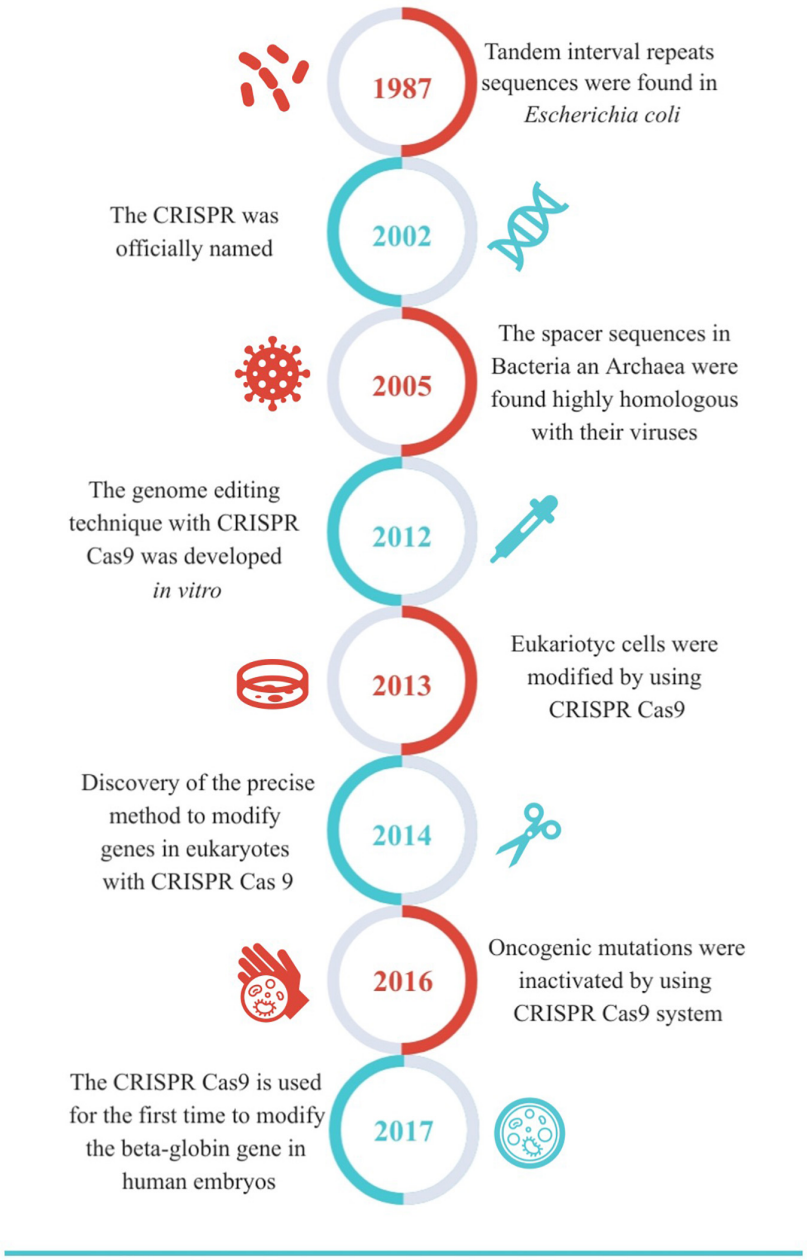


FIGURE 1 | Timeline of the most relevant discoveries about the CRISPR Cas system.

part. The bad part usually has the economic interests of large companies added (Wu et al., 2020).

Information on the application of the CRISPR Cas system should focus on deeply studying the damages or collateral effects of said system, when using it in living beings, and finally in humans. The feasibility of the target sites that are intended to be modified to solve, in general, a health problem, should be carefully studied in cell culture models, possibly also in animal models, waiting for favorable results, without the alteration of other metabolic factors. Derived from the

growing interest in the use and application of CRISPR Cas for genetic modification, the scientific community is often in ethical dilemmas, due to the division of ideas that, on the one hand, promise encouraging results, and on the other hand, there is the question of commitment of life in general (Cai and Wang, 2019; Pickar-Oliver and Gersbach, 2019; Wu et al., 2020).

Ethical dilemmas and the need for a body that guarantees the rights of living beings subjected to biomedical processes gave rise to bioethics, whose main task would be to ensure human rights and dictate the principles that ensure respect for life, in other words, stipulates the bases on which the investigations will be developed, how researchers should be governed and the bases of their investigations (Cribbs and Perera, 2017; Marinelli and Del Rio, 2020).

In 2014, Zhang's group and collaborators from the Broad Institute obtained a patent that granted them the right to use the CRISPR system in mice, humans, pigs, and almost any organism, other than bacteria. First controversial aspect on the use of this tool, is that the patent was obtained quickly, in less than six months. Furthermore, the works of Doudna and Charpentier, who had previously applied for a patent, had been rejected, as their possible use in humans had been speculated, contrary to what was published by Zhang, who had already tested it in humans, and had been in place since 2012 (Cong et al., 2013).

Nevertheless, in 2011, they had begun collaborative studies with the group of Doudna and Charpentier, who unified knowledge about *Streptococcus pyogenes* and RNA, respectively. The association of the two researchers allowed the beginning of the era of <genomic scissors>, together with the standardization of the *in vitro* method. The enzyme used for this purpose is the Cas 9 protein that acts on DNA through the guidance of RNA, forming a chimera. The simplification of the method allowed the conversion of a natural phenomenon into a genetic engineering technique. In such a way that this advance would allow the use of the system for gene editing in eukaryotic cells in a specific and precise way in predetermined sites, and not only providing immunity to bacteria (Doudna and Charpentier, 2014).

Later, Zhang's group, in 2017, gave a twist to the CRISPR Cas technology through its publication in which the ability to edit RNA by using the Cas 13 protein was presented, associating it with the adenosine deaminases protein (ADAR). The editing system was called REPAIR (RNA Editing for Programmable Adenosine to Inosine Replacement), this new technology would allow to change an adenosine base with a base inosine, to correct the point mutations that cause genetic diseases due to defects in the RNA. Subsequently, they sought to correct with an efficiency of 20–40%, experimenting with *in vitro* mutations of conditions, such as Fanconi anemia or nephrogenic diabetes insipidus, correcting them successfully by using the REPAIR system (D'Souza, 2017; Gootenberg et al., 2017).

Some variants of the CRISPR Cas system have been used for genome editing, due to the efficiency of gene editing and the wide scope of genome orientation, of which the Cas9 protein is the most widely used, of which, various investigations have focused in modifying the Cas9 protein and increasing its efficiency. The applications are diverse, epigenetic editing is listed for the specific alteration of loci, regulation of genes for the activation or deactivation of the expression of a gene or group of genes. The monitoring of cell dynamics by chromatin analysis in conjunction with the 3D modification of cell chromatin. By allowing DNA recognition and RNA editing, the application of the system is enormous, from biomedicine to biotechnology, some examples of the current CRISPR-Cas systems application related to control or to cure diseases are focused on multiple myeloma; esophageal, lung, prostate, and bladder cancer; solid tumors; melanoma; leukemia; human papilloma virus; HIV-1; gastrointestinal infection; β-thalassemia; sickle-cell anemia, among others (Brokowski and Adli, 2019).

Various investigations can be counted, however, each one of them has points that can be subjected to deep discussion by a scientific committee that evaluates and, where appropriate, approves or rejects them. This is due to the fact that the total effectiveness is still being observed and without adverse effects, that is, in the trial period. The doubt is that the Cas 9 protein has not shown a 100% effectiveness, since it has a relatively frequent variability of cut, so there is a diversity of modified sequences and some with some similarity to other mutations, which could generate adverse effects (Bhan et al., 2017; Caballero-Hernández et al., 2017).

Due to the lack of confidence in the CRISPR Cas system and its application in the cure of genetic diseases, voices have arisen claiming to be careful with releasing them, in addition to the fact that few or perhaps none of the countries are already thinking about laws that regulate the use of these techniques, in addition to the probable complication of patent application. The question arises as to whether the modified genetic sequences are patentable (Capella, 2016; Bhan et al., 2017).

The Limits of Genetic Modification

In the broad field of genetic engineering research, the number of laboratories with scientists dedicated to research this system has increased in recent years. Probably each one of them has proposed to solve a problem through this genetic technology, which will eventually fall into the patent fight. Moreover, the society demands the release of this technique to save lives worldwide, claiming that this discovery cannot be under a patent, since it would become almost or totally unattainable for many human beings; however, scientists warn about its release (Cribbs and Perera, 2017; Khan, 2019).

All over the world, genetic modification has been an object of reflection for bioethics, and this has been increased by the arrival, in large part, of the CRISPR Cas system. It is important to note that technology of CRISPR Cas is not inherently "good" or "bad," technology is tools and forms of power, which can be well illustrated by Michael Foucalt in his concept of biopolitics and the implementation of power over our bodies. In such a way that, the result of the use of new genetic modification technologies can be something applauded or something deplorable, taking into account that the tool used is not the one that determines the end, it is the user who determines the fact. This technology and the stem cell modification line represent a great potential for the development of revolutionary genetic therapies, representing a

feasible possibility of exploitation and clinical application. After the probable approval of CRISPR Cas as a therapeutic alternative, it is questioned how feasible it is to approve it, if it will be accessible to the public, in which cell lines it could be applied, in addition to the laws that should govern its use. The origin of the publications that have genetically modified human embryos by using CRISPR Cas 9, has caused different scientists to pronounce on the location or suspension of this type of research, however, these pronouncements should be accompanied by arguments on how to regulate these novel genetic tools (Brokowski, 2018; Lee and Kim, 2018).

When questioning to what extent the use of CRISPR in clinical medicine should be allowed, through the use of autonomy, it must clarify whether the user truly knows the risk of undergoing these treatments. Here, morality and what is legally permissible is considered, seeking to justify the use without considering the risk posed by research involving CRISPR-based genome engineering. Particularly, due to an important fact, the general risk profile of CRISPR experimentation in human beings remains unknown and it is the scientific duty to incur in these situations and to evaluate them objectively, eliminating dogmas, misperceptions, and personal prejudices, but always accompanied by institutional observation and with bioethical limits well established by specialized committees (Gil, 2019; Greely, 2019).

The great potential of genetic modification by means of CRISPR Cas, of cells or tissues, even of whole organisms, raises questions about its feasibility, since the modification does not remain only in the modified organism, but also in its offspring. The main argument of legalization can be divided into two main currents, the one that calls for laws that regulate the obtaining of patents to make these technologies accessible to patients who require it and those that regulate the use under controlled conditions, and the one where governments should be aware of who is developing studies about it. The latest due to the emergence in recent years of the so-called biohackers (Landrain, 2013; de Lecuona et al., 2017).

Biohackers

Biohacker communities have proliferated in the world, without anyone being able to stop them, who are dedicated to research, development, and innovation of all kinds of scientific and technological activities. These communities are dedicated to exceeding the limits of biology, arguing that pharmaceutical companies are enriched by the development of techniques that could be performed at home without any problem, without clinical control, and without medical supervision, in addition to the affordable cost. Its main task is to generate treatments to increase life, cure diseases, increase available treatments, and increase the physical, biological, and physiological capacities of humanity (Kera, 2012; Meyer, 2012; Landrain, 2013; Gil, 2019).

These communities support the fundamental ethical arguments, adding them to philosophical theories that had not even been touched in recent years. At the time of the alchemists, it was only intended to stop the deterioration of age, but nowadays the cure of genetic and motor diseases is sought, with the argument of increasing human capacity within the framework of freedom, through DIYBio or Do It Yourself Biology. The topic of DIYBio became relevant since 2005, when it was mentioned at CodeCon, the possibility of purifying DNA at home with simple objects led to the promotion of free research in DNA biology and, together with various ethical controversies on the synthetic biology, the modification of eukaryotic organisms, with the problem of having information on biological techniques freely accessible, without legal regulation, control, and validation of a scientific organism (Meyer, 2012; Landrain, 2013; Gil, 2019).

In the United States Federal Drug Administration (FDA) has detected the presence of biohackers in its territory, they have experimented with various treatments from their own garage, highlighting the intradermal injection of DNA molecules modified by CRISPR Cas, that promise to cure a disease. The persecution of biohackers in the USA has been given for the practice of medicine without a license, however, these individuals could be accused of misuse of medical treatment, but to date in North American laws, the CRISPR Cas is not considered medical treatment, for which these acts could not be condemned, which results in the importance of countries, including the World Health Organization, of the guidelines for the international regulation of the use of CRISPR Cas as medical treatment and who, how, and where these biological technologies can be developed (Delfanti, 2011; Hirsch et al., 2019).

CRISPR-Cas MODIFICATION, LEGAL REGULATION

Since 1975, at the Asilomar conference, a growing concern was expressed about the use of recombinant DNA, despite the fact that the use of technology applied to DNA was allowed, the bioethical arguments regarding the application of genetic engineering to humanity continues to be the subject of deep debate (Evitt et al., 2015; Capella, 2016; Brokowski and Adli, 2019). In the 1960s, the theory of gene therapy in Stanfield's experimental trial, referring to congenital metabolic diseases, was questioned regarding the ethical problems surrounding its execution. Until 1990, gene therapy was approved in humans, at least at the subclinical level, in such a way that it was confirmed in the Whiley Database on Gene Therapy Clinical Trials Worldwide, of the National Institute of Health (NIH) of the United States. However, until 2000 there were indications of the reliability of these therapies in humans, after arduous attempts to improve the technique for the treatment of severe combined adenosine deaminase immunodeficiency (ADA-SCID) (Espinosa and Hernández-Hernández, 2020).

Genetic modification is a well-known topic for bioethics, which is far from being forgotten and perhaps further from being resolved. The first issue in question is whether to allow the use of CRISPR Cas technology for gene modification, since it is doubtful that it can be put into practice in humans. Various statements around the world have expressed their concern about the regulation of gene editing. In the United Kingdom in 2015, the meeting of the Hinxton Group and the international meeting of the Academy of Sciences of the United States of North America took place, arguing that for no reason should clinical research applied to the clinic be limited, due to the concern of misuse.

With the limitations imposed by bioethics in gene editing in embryos, the experiments were allowed by the Authority for Human Embryology and Fertilization of the United Kingdom, with the controversy crossed over the generation of humans modified by CRISPR Cas with resistance to HIV (Capella, 2016; Gamboa-Bernal, 2016; Hirsch et al., 2019).

Currently, Mexico seems to be a cradle for the development of stem cell treatments; nevertheless, the growth in the availability of these treatments makes Mexico a destination susceptible to bioethical conflicts due to the relatively easy obtaining of unproven applications, given the lack of scientific and legal regulation. Research with humans or tissues from humans in Mexico is governed by the General Health Law and the Regulation of the Law on Research Matters, it includes the Official Mexican Standard NOM-012-SSA3-2012, and the Declaration of Helsinki, which together provide the statutes by which the institutional Bioethics committees analyze the approval or prohibition of preclinical and clinical studies where humans are studied. The restrictions and limitations that researchers will have in this matter are listed in international treaties (López-Pacheco et al., 2016; Espinosa and Hernández-Hernández, 2020).

The genetic modification techniques based on CRISPR Cas are extremely novel, due to their qualities and relative ease of design, but not of execution, which allows the genetic modification of humans. The arguments that are presented against its use are valid, due to the fact of the possible alteration of the physiology of the organism, without ceasing to consider it beneficial if a health problem is corrected. Havoc or damage could translate into long-term damage that, translated into evolutionary events, would be a catastrophic event if not taken seriously. Alternatively, the concern of the scientific community focuses on the fact that there are no real limits for scientists, from a legal point of view, therefore, it is not clear how far the power of genetic modification can go and the weight that has on bioethical practices that seek to always go toward respecting life and the rights of living beings, in addition to the implementation of the four principles of bioethics. The scrutiny of current knowledge of these technologies, the way that these can either help or fail to achieve desired modifications, and the future promise and challenge of therapeutic genome editing, should be open for discussion, not only by scientists and physicians, in order to overcome the problem of Techno-Scientific Colonialist Paternalism (Arguedas-Ramírez, 2020).

We should consider the insight given by bioethicists in this subject, to mediate and be aware of the pros and cons of these new technologies, highlighting the importance of an open public discussion in which both parties are taken into consideration: the scientific and ethical facts to define the real issues in this picture. The role of the government in these regulations and instances to be addressed, should be to regulate and make sure to satisfy the needs to benefit all the society in need of these technologies, not only the privileged population. The latter, to ensure the ethical use of these systems to try to reduce gaps and social inequalities instead of opening new ones, while considering the current sociopolitical, economical, and historical issues, such as the anti-scientific movements and the politicization of science (Arguedas-Ramírez, 2020).

Therefore, the content of the current paper relies on opening a discussion in which the currently known real issues surrounding these technologies are described, in order to overcome the fear and doubts that can be generated in the global population.

Future Perspectives of CRISPR Cas

The use of CRISPR Cas has generated an enormous progress in the development of biotechnology, it is the most outstanding discovery of the 21st century. The vision for the future that is expected from the CRISPR Cas system, from an anthropocentric position, is the cure of rare or catastrophic diseases such as cancer, diabetes, or congenital anemia, among others. In the health area, the application of this system would allow the fight against diseases such as HIV, malaria, dengue, Zika virus or the current SARS-CoV-2. In addition to the above, there is the possibility of interrupting the advance of bacterial resistance to antibiotics, in addition to decreasing the virulence of bacterial isolates that cause infections. The purpose is to always try to help human health, but care must be taken not to cause unwanted alterations in the patient (Escalona-Noguero et al., 2021; Tsou et al., 2021; Yadav et al., 2021).

However, the area of application of this system is enormous, agriculture should benefit from the CRISPR Cas helping to improve food and increasing availability to combat famine and food shortages. In the biotechnology industry, the CRISPR Cas suggests being valuable for the implementation or modification of metabolic routes, this would increase the yields in obtaining the product, optimizing the processes for obtaining products of biological origin (Nidhi et al., 2021).

The use of CRISPR Cas implies legal and bioethical principles, initially these principles should protect human dignity, safeguard the integrity of the patient, and safeguard the content of their genetic information to avoid inappropriate uses. Another aspect that should be considered is solving the inherent risk of undergoing these treatments that will not show adverse effects, if any, in the short-term. Together, bioethics and legal law will have to work together to regulate the use of patient information, protecting fundamental rights, such as health. Considering that these techniques must be carefully evaluated and observed, not to prohibit them, but to handle them with care, because it is known that with this technique some diseases could be eradicated (Nidhi et al., 2021).

AUTHOR CONTRIBUTIONS

LG-A, CH-C, and GC-E developed the structural design of the review and drafted the manuscript. LG-A, JV-L, LP-R, OC-M, CH-C, and GC-E reviewed the manuscript critically for important intellectual content and appropriate academic content. All authors read and approved the final manuscript.

FUNDING

LG-A and OC-M received grant-aided support from "Consejo Nacional de Ciencia y Tecnología" (CONACyT, México), while CH-C and GC-E received support from the "Sistema Nacional de Investigadores (SNI)" from CONACyT, Mexico. GC-E received support from Estímulos al Desempeño en Investigación,

Comisión y Fomento de Actividades Académicas (Instituto Politécnico Nacional). This study was funded by Secretaría de Investigación y Posgrado del Instituto Politécnico Nacional (SIP 20200675, 20211351).

ACKNOWLEDGMENTS

The authors are grateful to Sofia Mulia for correcting the style of the manuscript.

REFERENCES

Arguedas-Ramírez, G. (2020). Ethics and Global Governance of Human Germline Genome Editing: The Problem of Techno-Scientific Colonialist Paternalism. *CRISPR J.* 3, 83–88. doi: 10.1089/crispr.2019.0045

Barrangou, R., and Marraffini, L. A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54, 234–244. doi: 10.1016/j.molcel.2014.03.011

Bhan, A., Sasikumar, S., Goja, A., and Tk, R. (2017). Bioethical Concerns of CRISPR: a genome editing technology. *Glob. Bioeth. Enquiry* 13:68. doi: 10.38020/GBE.5.2.2017.68-72

Brokowski, C. (2018). Do CRISPR Germline Ethics Statements Cut It? *CRISPR J.* 1, 115–125. doi: 10.1089/crispr.2017.0024

Brokowski, C., and Adli, M. (2019). CRISPR Ethics: Moral Considerations for Applications of a Powerful Tool. *J. Mol. Biol.* 431, 88–101. doi: 10.1016/j.jmb.2018.05.044

Caballero-Hernández, D., Rodríguez-Padilla, C., and Lozano-Muñiz, S. (2017). Bioethics for biotechnologists: from dolly to CRISPR. *Open Agricult.* 2, 160–165. doi: 10.1515/opag-2017-0016

Cai, W., and Wang, M. (2019). Engineering nucleic acid chemistry for precise and controllable CRISPR/Cas9 genome editing. *Sci. Bull.* 64, 1841–1849. doi: 10.1016 /j.scib. 2019.07.035

Capella, V. B. (2016). La revolución de la edición genética mediante CRISPR-Cas 9 y los desafíos éticos y regulatorios que comporta [The CRISPR-Cas 9 gene editing revolution and the ethical and regulatory challenges it brings]. *Cuadernos Bioét.* 27, 223–239.

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

Cribbs, A. P., and Perera, S. (2017). Science and Bioethics of CRISPR-Cas9 Gene Editing: An Analysis Towards Separating Facts and Fiction. *Yale J. Biol. Med.* 90, 625–634.

D'Souza, R. (2017). Ethical Issues associated with diagnosing and managing diabetes. *Glob. Bioeth. Enquiry* 2017, 52–56. doi: 10.38020/gbe.5.2.2017.52-56

de Lecuona, I., Casado, M., Marfany, G., Lopez Baroni, M., and Escarrabill, M. (2017). Gene Editing in Humans: Towards a Global and Inclusive Debate for Responsible Research. *Yale J. Biol. Med.* 90, 673–681.

Delfanti, A. (2011). *Genome hackers. Rebel biology, open source and science ethic*. Milano: Università degli studi di Milano, doi: 10.13130/delfanti-alessandro_phd2011-07-14

Doudna, J. A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1258096.

Escalona-Noguero, C., López-Valls, M., and Sot, B. (2021). CRISPR/Cas technology as a promising weapon to combat viral infections. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 43:e2000315. doi: 10.1002/bies.202000315

Espinosa, A. R. B., and Hernández-Hernández, M. (2020). La bioética del sistema CRISPR-Cas9 como terapia génica en enfermedades de importancia mundial [The bioethics of the CRISPR-Cas9 system as gene therapy in diseases of global importance]. *Ciencia Huasteca Boletín Científico de la Escuela Superior de Huejutla* 8, 29–33. doi: 10.29057/esh.v8i16.5674

Evitt, N. H., Mascharak, S., and Altman, R. B. (2015). Human Germline CRISPR-Cas Modification: Toward a Regulatory Framework. *Am. J. Bioeth. AJOB* 15, 25–29. doi: 10.1080/15265161.2015.1104160

Faure, G., Makarova, K. S., and Koonin, E. V. (2019). CRISPR-Cas: Complex Functional Networks and Multiple Roles beyond Adaptive Immunity. *J. Mol. Biol.* 431, 3–20. doi: 10.1016/j.jmb.2018.08.030

Gamboa-Bernal, G. A. (2016). La edición de Genes a Estudio: Los Problemas Bioéticos que puede tener esta nueva tecnología [The edition of Genes to Study: The Bioethical Problems that this new technology may have]. *Persona Bioética* 20, 125–131. doi: 10.5294/pebi.2016.20.2.1

Gil, L. F. R. (2019). Bioética y contracultura: el Biopunk y ética DIYBio [Bioethics and counterculture: Biopunk and DIYBio ethics]. *Rev. Colomb. Bioética* 14:2443.

Gómez-Tatay, L., and Aznar, J. (2019). CRISPR-Cas9. El mayor avance en técnicas de edición genética requiere una reflexión Ética [CRISPR-Cas9. The greatest advance in gene editing techniques requires ethical reflection]. *Cuaderno Bioética* 30, 171–185. doi: 10.30444/CB.31

Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Essletzbichler, P., Dy, A. J., Joung, J., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356, 438–442. doi: 10.1126/science.aam9321

Greely, H. T. (2019). Human Germline Genome Editing: An Assessment. *CRISPR J.* 2, 253–265. doi: 10.1089/crispr.2019.0038

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Hirsch, F., Iphofen, R., and Koporc, Z. (2019). Ethics assessment in research proposals adopting CRISPR technology. *Biochem. Med.* 29:020202. doi: 10.11613/BM.2019.020202

Ishino, Y., Krupovic, M., and Forterre, P. (2018). History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *J. Bacteriol.* 200, 580–517e. doi: 10.1128/JB.00580-17

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi: 10.1128/jb.169.12.5429-5433.1987

Jamal, L., Schupmann, W., and Berkman, B. E. (2020). An ethical framework for genetic counseling in the genomic era. *J. Genet. Counsel.* 29, 718–727. doi: 10.1002/jgc4.1207

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

Kera, D. (2012). Hackerspaces and DIYbio in Asia: connecting science and community with open data, kits and protocols. *J. Peer Product.* 2, 1–8.

Khan, S. H. (2019). Genome-editing technologies: concept, pros, and cons of various genome-editing techniques and bioethical concerns for clinical application. *Mol. Therapy Nucleic Acids* 16, 326–334. doi: 10.1016/j.omtn.2019.02.027

Landrain, T. (2013). Le Do-it-yourself biology: défis et promesses [Do-it-yourself biology: challenges and promises]. *Med. Sci.* 29, 33–35. doi: 10.1051/medsci/201329s209

Lee, H., and Kim, J. S. (2018). Unexpected CRISPR on-target effects. *Nat. Biotechnol.* 36, 703–704. doi: 10.1038/nbt.4207

López-Pacheco, M. C., Pimentel-Hernández, C., Rivas-Mirelles, E., and Arredondo-García, J. L. (2016). Normatividad que rige la investigación clínica en seres humanos y requisitos que debe cumplir un centro de investigación para participar en un estudio clínico en México [Regulations governing clinical research in humans and requirements that a research center must meet to participate in a clinical study in Mexico]. *Acta Pediátr. México* 37, 175–182. doi: 10.18233/apm37no3pp175-182

Makarova, K. S., Wolf, I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S., et al. (2020). 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., 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

Marinelli, S., and Del Rio, A. (2020). Beginning of life ethics at the dawn of a new era of genome editing: are bioethical precepts and fast-evolving biotechnologies irreconcilable? *La Clin. Terapeut.* 171, 407–411e.

Marraffini, L. A., and Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190. doi: 10.1038/nrg2749

Meyer, M. (2012). Build your own lab: Do-it-yourself biology and the rise of citizen biotech-economies. *J. Peer Product.* 2:710829.

Mojica, F. J., and Rodriguez-Valera, F. (2016). The discovery of CRISPR in archaea and bacteria. *FEBS J.* 283, 3162–3169. doi: 10.1111/febs.13766

Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evolut.* 60, 174–182. doi: 10.1007/s00239-004-0046-3

Nidhi, S., Anand, U., Oleksak, P., Tripathi, P., Lal, J. A., Thomas, G., et al. (2021). Novel CRISPR-Cas Systems: An Updated Review of the Current Achievements, Applications, and Future Research Perspectives. *Int. J. Mol. Sci.* 22:3327. doi: 10.3390/ijms22073327

Noll, S. E. (2019). Review of Genetic Ethics: An Introduction, by Colin Farrelly. *Essays Philosop.* 20, 245–250. doi: 10.7710/1526-0569.1638

Pérez, M. G. (2006). Los principios de la bioética y la inserción social de la práctica médica [The principles of bioethics and the social insertion of medical practice]. *Rev. Administ. Sanitaria* 4, 341–356.

Pickar-Oliver, A., and Gersbach, C. A. (2019). The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* 20, 490–507. doi: 10.1038/s41580-019-0131-5

Schupmann, W., Jamal, L., and Berkman, B. E. (2020). Re-examining the Ethics of Genetic Counselling in the Genomic Era. *J. Bioeth. Inquiry* 17, 325–335. doi: 10.1007/s11673-020-09983-w

Shkomova, E. (2019). Transformation of the ethical Regulations in the Context of the medical Genetic Counseling Improvement. *Chelovek* 30, 145–155. doi: 10.31857/S023620070007677-1

Tsou, J. H., Liu, H., Stass, S. A., and Jiang, F. (2021). Rapid and Sensitive Detection of SARS-CoV-2 Using Clustered Regularly Interspaced Short Palindromic Repeats. *Biomedicines* 9:239. doi: 10.3390/biomedicines9030239

Wu, S. S., Li, Q. C., Yin, C. Q., Xue, W., and Song, C. Q. (2020). Advances in CRISPR/Cas-based Gene Therapy in Human Genetic Diseases. *Theranostics* 10, 4374–4382. doi: 10.7150/thno.43360

Yabar, M. L. P. (2019). CRISPR/Cas9: Edición Genética vs. Bioética: Un análisis bioético de la técnica de edición de genes llamada CRISPR/Cas9 [CRISPR / Cas9: Genetic Editing vs. Bioethics: A bioethical analysis of the gene editing technique called CRISPR/Cas9]. *Apuntes Bioét.* 2, 5–18.

Yadav, N., Narang, J., Chhillar, A. K., and Rana, J. S. (2021). CRISPR: A new paradigm of theranostics. *Nanomed. Nanotechnol. Biol. Med.* 33:102350. doi: 10.1016/j.nano.2020.102350

You, L., Tong, R., Li, M., Liu, Y., Xue, J., and Lu, Y. (2019). Advancements and Obstacles of CRISPR-Cas9 Technology in Translational Research. *Mol. Therapy Methods Clin. Dev.* 13, 359–370. doi: 10.1016/j.omtm.2019.02.008

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 02 June 2021
doi: 10.3389/fmicb.2021.656996

Identification and Characterization of the CRISPR/Cas System in *Staphylococcus aureus* Strains From Diverse Sources

*Erick Adrian Cruz-López[1], Gildardo Rivera[1], María Antonia Cruz-Hernández[1], Ana Verónica Martínez-Vázquez[1], Graciela Castro-Escarpulli[2], Rebeca Flores-Magallón[3], Karina Vázquez[4], Wendy Lizeth Cruz-Pulido[5] and Virgilio Bocanegra-García[1]**

[1]*Laboratorio Interacción Ambiente-Microorganismo, Centro de Biotecnología Genómica, Instituto Politécnico Nacional, Tamaulipas, Mexico,* [2]*Laboratorio de Investigación Clínica y Ambiental, Departamento de Microbiología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Mexico City, Mexico,* [3]*Centro Interdisciplinario de Investigación Para el Desarrollo Integral Regional, Unidad Michoacán, Instituto Politécnico Nacional, Jiquilpan, Mexico,* [4]*Facultad de Medicina Veterinaria, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico,* [5]*Research Department, Universidad del Valle de México, Campus Reynosa, Mexico City, Mexico*

The CRISPR-Cas [clustered regularly interspaced short palindromic repeats and the CRISPR-associated genes (Cas)] system provides defense mechanisms in bacteria and archaea vs. mobile genetic elements (MGEs), such as plasmids and bacteriophages, which can either be harmful or add sequences that can provide virulence or antibiotic resistance. *Staphylococcus aureus* is a Gram-positive bacterium that could be the etiological agent of important soft tissue infections that can lead to bacteremia and sepsis. The role of the CRISPR-Cas system in *S. aureus* is not completely understood since there is a lack of knowledge about it. We analyzed 716 genomes and 1 genomic island from GENOMES-NCBI and ENA-EMBL searching for the CRISPR-Cas systems and their spacer sequences (SSs). Our bioinformatic analysis shows that only 0.83% (6/716) of the analyzed genomes harbored the CRISPR-Cas system, all of them were subtype III-A, which is characterized by the presence of the *cas10/csm1* gene. Analysis of SSs showed that 91% (40/44) had no match to annotated MGEs and 9% of SSs corresponded to plasmids and bacteriophages, indicating that those phages had infected those *S. aureus* strains. Some of those phages have been proposed as an alternative therapy in biofilm-forming or infection with *S. aureus* strains, but these findings indicate that such antibiotic phage strategy would be ineffective. More research about the CRISPR/Cas system is necessary for a bigger number of *S. aureus* strains from different sources, so additional features can be studied.

Keywords: *Staphylococcus aureus*, methicillin-resistant *Staphylococcus aureus*, multidrug resistant, CRISPR-Cas system, phage therapy

OPEN ACCESS

Edited by:
Friedrich Götz, University of Tübingen, Germany

Reviewed by:
Débora Coraça-Huber, Innsbruck Medical University, Austria
Ananda Shankar Bhattacharjee, University of Illinois at Urbana-Champaign, United States

****Correspondence:***
Virgilio Bocanegra García
vbocanegg@hotmail.com;
vbocanegra@ipn.mx

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *21 January 2021*
Accepted: *05 May 2021*
Published: *02 June 2021*

Citation:
Cruz-López EA, Rivera G, Cruz-Hernández MA, Martínez-Vázquez AV, Castro-Escarpulli G, Flores-Magallón R, Vázque K, Cruz-Pulido WL and Bocanegra-Garcia V (2021) Identification and Characterization of the CRISPR/Cas System in Staphylococcus aureus Strains From Diverse Sources. Front. Microbiol. 12:656996. doi: 10.3389/fmicb.2021.656996

INTRODUCTION

The bacteria and archaea have developed defense mechanisms against bacteriophages (Labrie et al., 2010), in the form of restriction and modification system (R-M *system*; Huff et al., 2017) and as the CRISPR-Cas [clustered regularly interspaced short palindromic repeats and the CRISPR-associated genes (Cas)] system, both of which degrade the foreign invader genetic material. The CRISPR-Cas is a natural, memory, and hereditary mechanism that protects bacteria against bacteriophages (Faruque et al., 2005; Box et al., 2015; Hille et al., 2018). It is composed of (1) a group of genes *cas*, (2) a locus or loci, CRISPR formed by spacer sequences (SSs) separated into repeated sequences (SRs), and (3) the leader sequence placed upstream from locus CRISPR (Westra et al., 2014); the set of *cas* genes is divided into the module of adaptation formed by *cas1* and *cas2* genes and the effector complex where the rest of the *cas* genes are placed (Koonin et al., 2017).

The system CRISPR-Cas current classification includes 2 classes, 6 types, and 33 subtypes. Class 1 systems use multi-unit protein complexes (Koonin and Makarova, 2017; Koonin et al., 2017) and Class 2 systems use only one multidomain protein (Shmakov et al., 2017) for the degradation of the genetic material. This DNA degradation occurs in three stages (Hsu et al., 2014): (1) adaptation stage during a primo-infection (Nuñez et al., 2014), (2) expression stage during reinfection, and (3) interference stage for the digestion mobile genetic element (MGE) through endonucleases Cas, which are guided by crRNA (chimera of SE and SR; Hille et al., 2018). In a MGE, there are short sequences (approximately 30 nucleotides) marked by protospacer adjacent motifs (Jiang and Doudna, 2015, 2017), known as protospacers, which are inserted like a SS.

The CRISPR-Cas system has been detected in Gram-positive bacteria, such as *Lactobacillus* spp. (Wang et al., 2020; Yang et al., 2020) and pathogenic bacteria, such as *Enterococcus* spp. (Sanderson et al., 2020). However, in *Staphylococcus aureus*, it has only been detected in few strains. *S. aureus* is a Gram-positive bacterium that colonizes 30% of the population in an asymptomatic way, and also it is the etiological agent of several important infections (Craft et al., 2019). In 1960, the methicillin-resistant *S. aureus* (MRSA) strains were detected (Chambers and Deleo, 2009), and those are still a current cause of soft tissue infections. New effective antibiotic therapies are a current demand (Yang et al., 2018; Labruère et al., 2019). The application of bacteriophages as a therapy to treat *S. aureus* infections (Kaźmierczak et al., 2014) is a promising alternative. The memory capacity of the CRISPR-Cas system allows us to understand the dynamic between an MGE and its hosts (bacteria and archaea). The sequenced bacterial genomes are the current data source for searching CRISPR-Cas system in important medical bacteria such as *S. aureus*. Despite studies searching CRISPR-Cas system in *S. aureus* (Cao et al., 2016; Zhao et al., 2018; Rossi et al., 2019), further research is needed to study this system in more *S. aureus* strains to understand the effects and its association to pathogenicity. Thus, the aim of this study was to search CRISPR-Cas in *S. aureus* genomes and its characterization *via* bioinformatic tools, as well as the association of the SS with MGEs.

MATERIALS AND METHODS

Genomes Collection

The complete *S. aureus* genomes were downloaded from GENOME-NCBI [n = 864 (484 chromosomes and 380 plasmids)] and ENA-EMBL [n = 521 (232 chromosomes, 288 plasmids, and 1 pathogenicity island)]. In total, there were 716 strains used.

CRISPR-Cas System Determination

The genomes were analyzed using CRISPRCasFinder 4.2.2 (Grissa et al., 2007; Abby et al., 2014; Couvin et al., 2018). The server was used with default parameters: minimal repeat length 23 bp, maximal repeat length 55 bp, repeat mismatch, minimal spacer size in function of repeat size 0.6, maximal spacer size in function of repeat size 2.5, maximally allowed percentage of similarity between spacers 60, percentage mismatches allowed between repeats 20, percentage mismatches allowed for truncated repeat 33. Also, a default 100 bp size of flanking regions in all potential CRISPR arrays was included. A CRISPR-Cas system that presents a group of genes *cas* and the locus CRISPR with a score of 3 and 4 were considered for the next analysis.

Cas1, Cas2, Cas6, and Cas10 Phylogenetic Analysis Proteins

The files containing the coding sequences of each CRISPR-Cas system-bearing genome were downloaded from GENOME-NCBI. The *cas* genes were obtained from those files and translated into MEGA-X by using the standard genetic code. Later, the Cas proteins were aligned to the program Clustal O of Unipro UGENE. The scores "pair sum" were calculated in GeneDoc. The best alignments showed lower scores. Subsequently, the phylogenetic trees were built by the UPGMA method using default parameters with 1,000 bootstrap in the program MEGA-X (Kumar et al., 2018).

Cas Protein Analysis

The Cas sequences annotated images were generated in the program EasyFig 2.2.5 (Sullivan et al., 2011).

Phylogenetic Analysis of the Repeated Sequences

The analysis was the same process as the Cas proteins, except that we used the neighbor-joining method with default parameters to build the phylogenetic tree, using MEGAX (Rose et al., 2019).

Secondary Structures of Repeat Sequence Consensus

The secondary structures of repeat sequence consensus (SRc) and the minimum free energy (MFE) were obtained from the RNAfold web server (http://rna.tbi.univie.ac.at//cgi-bin/

RNAWebSuite/RNAfold.cgi; Lorenz et al., 2011). The logo of SRc was obtained from WebLogo (Schneider and Stephens, 1990; Crooks et al., 2004).

Spacer Sequence Analysis

The FASTA files were downloaded from CRISPR-CasFinder. Spacer sequence (SE) were submitted to BLAST, and the results associated with the MGE were the ones considered with expected values (*e*-values) minor or equal to 0.0001 and scores above 40 (Ostria-Hernández et al., 2015). Next, a 0 and 1 matrix was developed, 1 being the cell where the MGE and *S. aureus* strain intercept. That matrix was analyzed in the ClustVis server using default parameters. Then, a heat map was elaborated with the webserver ClustVis (Metsalu and Vilo, 2015), where the MGE known as the aforementioned infected strain of *S. aureus* was appreciated.

RESULTS

The CRISPR-Cas system was searched in 1,385 sequences of *S. aureus*, including chromosomes, plasmids, and 1 pathogenicity island, collected from 2 databases (**Supplementary Table A**). The search showed that 0.83% (6/716) of *S. aureus* strains harbored the CRISPR-Cas system. The strains harboring the CRISPR-Cas system were *S. aureus* 08BA02176 (NC_018608), *S. aureus* KUH140087 (NZ_AP020315), *S. aureus* JS395 (NZ_CP012756), *S. aureus* AR_0470 (NZ_CP029653), *S. aureus* AR_0472 (NZ_CP029649), and *S. aureus* AR_0473 (NZ_CP029681). All these strains have different geographical origin: *S. aureus* 08BA02176 was isolated from a surgery infection in 2008 from a Canadian patient (Golding et al., 2012); *S. aureus* KUH140087 was isolated in 2014 from a septicemia patient in Kyoto, Japan (Hikichi et al., 2019); *S. aureus* JS395 was isolated in 2008 from a patient in Switzerland (Larsen et al., 2017), and the *S. aureus* strains AR_0472, AR_0470, and AR_0473 were sent by the Center for Disease Control and Prevention. While the source of *S. aureus* AR_0472, AR_0470, and AR_0473 is uncertain, the rest of the strains come from clinic sources. All the detected CRISPR-Cas systems were found in chromosome structures, and none were detected in the pathogenicity island; nevertheless, other islands had it (Chakraborty et al., 2009; Carpenter et al., 2017). The detected systems found belong to the III-A subtype, which is characterized by the gen *cas10/csm1* and *cas* genes ordered as shown in **Figure 1** (Koonin et al., 2017). The detected CRISPR-Cas system structure was as follows: (1) *cas* genes nearby the locus CRISPR and (2) scores of 3 and 4 in the CRISRPRCasFinder scale (Pourcel et al., 2020; **Supplementary Table B**). The strain contains a group of *cas* genes (**Figure 1**) near the locus CRISPR-Cas; the CRISPR locus and the cluster *cas* are separated by 73 nt (strains JS395 and AR_0470), 74 nt (strains 08BA02176, AR_0472, and AR_0473), and 133 nt (strain KUH140087). Each strain with CRISPR-Cas has a unique locus CRISPR with a different number of SS.

The SRc was the same in two strains (i.e., AR_0472 and AR_0473) and different in four strains. The SRc length was 36 and 37 nt; the SRc formed by 37 nt is shown in the strains JS395 and AR_0479. The SRc remains nucleotide motifs (**Figure 2**) that stand out among the conservative nucleotides: four consecutive nucleotides of cysteine (-CCCC-) and four consecutive nucleotides of guanine (-GGGG-). Among the conserved motifs, there is a constant region of eight nucleotides.

The Cas proteins and SR keep a function–structure relationship (*stem-loop* structure); the coevolution of both structures is necessary for the correct function of the system CRISPR-Cas. **Figures 3A**,**B** show the phylogenetic relations of the Cas proteins and SRc, respectively. The Cas proteins and SRc present in *S. aureus* KUH140087 are phylogenetically away from the ones present in the phylogroup formed by the rest of the strains.

The SR keeps the nucleotides that form the *stem-loop* structure, which gives the signals of the location where the cuts must be done on pre-crRNA. **Figure 4** shows the SRc secondary structures of the CRISPR-Cas systems found; the *stem* formed by interactions G:C (guanine:cysteine) can be seen, and the *loop* also indicates the MFE of each structure.

The memory and adaptation characteristics of the CRISPR-Cas system allow the bacteria to identify which MGE infected it. The subtraction of the protospacer from the MGE and its incorporation as SS in the locus CRISPR during the adaptation phase (McGinn and Marraffini, 2019) becomes an advantage in the genomic analysis. The total of SS [6 SS (strain JS395), 12 SS (strain AR_0470, AR_0472, and AR_0473), and 15 SS (strain 08BA02176)] is 62, where 26 (41.93%) are unique SS and 18 are SS duplicated and built 58.07%. Interestingly, the duplicated SSs are in the loci of strains AR_0472, AR_0473, JS395, and AR_0470. The SSs are preserved between the loci CRISPR: the repeated SSs of the strain AR_0472 (n = 12) match in order and sequence with the SS of the strain AR_0473 (n = 12), and the loci CRISPR of the strain JS395 (n = 6) also matches in order and sequence with a final 50% of loci CRISPR of strain AR_0470 (n = 12). Hence, only 44 SSs were considered for BLAST analysis.

The BLAST analysis showed that 9% (4/44) of SSs match with known MGE (**Supplementary Table C**). In **Figure 5**, MGEs that infected the *S. aureus* 08BA02176 (SS6) and *S. aureus* KUH140087 (SS1, SS2, and SS3) are presented. However, if a SS is associated with more than one MGE, it means that the protospacers are conserved between MGEs. Besides, according to our analysis, a specific protospacer can be found in plasmids

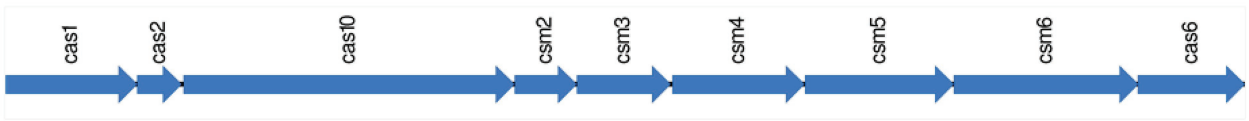


FIGURE 1 | The cluster of genes *cas* of the CRISPR-Cas subtype III-A of *S. aureus* system. The annotation genes *cas* was done in EasyFig 2.2.5.

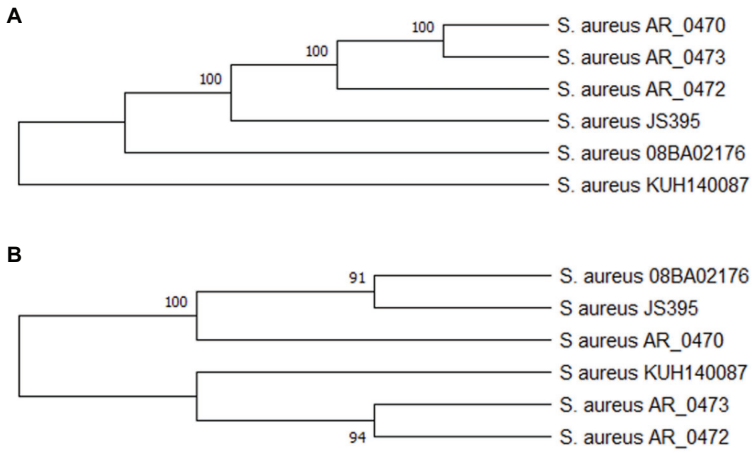


FIGURE 2 | Protein Cas phylogenetic relation **(A)** and SRc **(B)**. The alignment of the amino acid sequences of the proteins Cas and SRc was done with ClustalO, and the phylogenetic trees of SRc and the Cas proteins were built with the neighbor-joining and UPGMA method. There is a tree for the protein Cas (Cas1, Cas2, Cas10/Csm1, and Cas6) because they presented the same phylogenetic relation. The trees are the consensus of 1,000 bootstrap, and they were done with MEGA X.

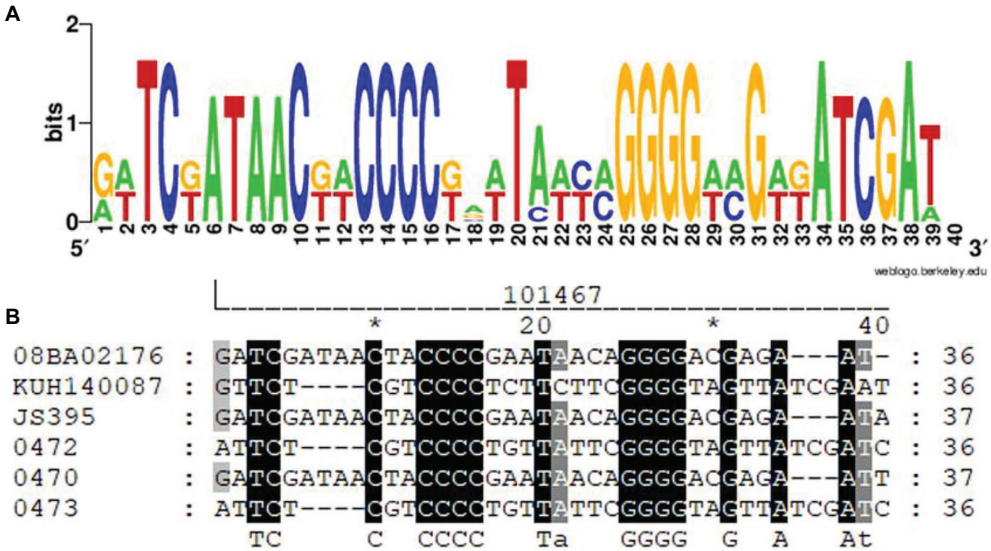


FIGURE 3 | SRc alignment of the CRISPR-Cas system present in *S. aureus*. Visualization **(A)** and alignment of nucleotides **(B)**. The motive nucleotides are under the alignment (capital letters). The alignment was done with MUSCLE (UGENE), and the image was obtained from WebLogo **(A)** and GeneDoc **(B)**.

or bacteriophages, but a plasmid protospacer is not found in a bacteriophage and vice versa. The SS1 of *S. aureus* KUH140087 is the only one that interferes with the two plasmids named in **Figure 5**, and the rest of SS interferes with the bacteriophage.

DISCUSSION

The CRISPR-Cas system is a heritable mechanism of immunity in bacteria and archaea, which protects them from foreign plasmids and bacteriophages; it is an endonuclease mechanism guided by crRNA (Makarova et al., 2013). Few studies have searched for the CRISPR-Cas system in *Staphylococcus* spp., where the CRISPR-Cas system was found in 0.94% (6/616) of isolated clinics (Cao et al., 2016) and in 7.89% (3/39) of the *S. aureus* strains analyzed for Zhao et al. (2018); moreover, the CRISPR-Cas system was searched in 129 isolated from *Staphylococcus* spp. (*S. aureus n* = 53, *Staphylococcus pseudintermedius n* = 74, *Staphylococcus haemolyticus n* = 1, and *Staphylococcus cohnii n* = 1) from 9 countries, and the 8% (10/129) are CRISPR-Cas system-bearing strains, but it was detected only in *S. pseudintermedius* strains (Rossi et al., 2019). Few studies have searched the CRISPR-Cas system in MGE, such as plasmids (Kamruzzaman and Iredell, 2020) or

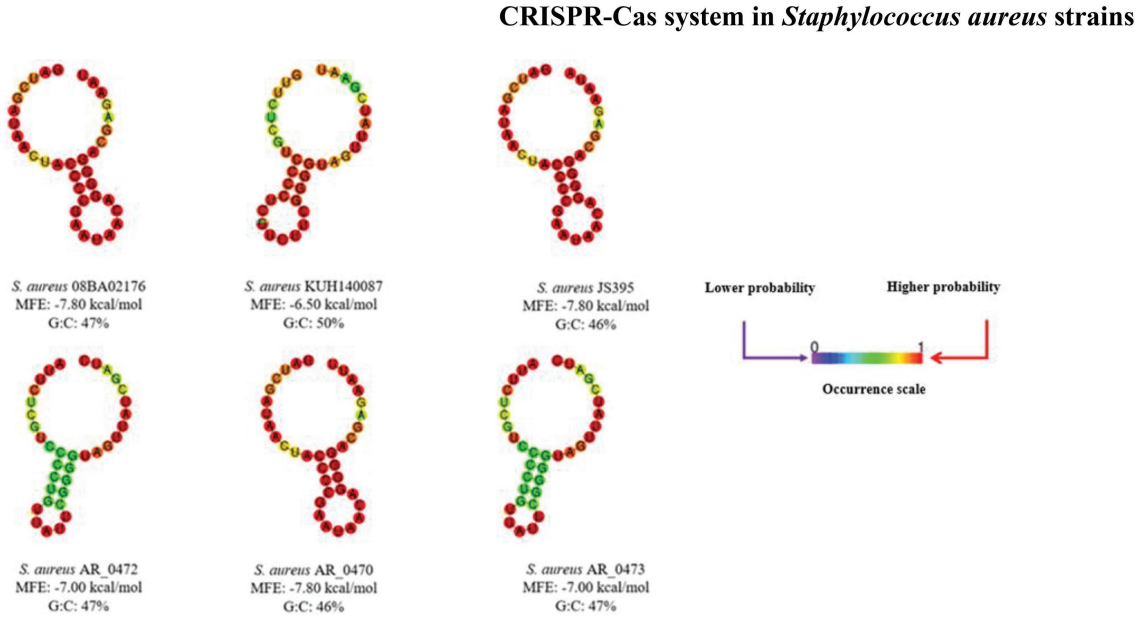


FIGURE 4 | Secondary structures of repeated sequences. Each secondary structure is the result of the interactions of the nucleotides; these structures and the minimum free energy were obtained in the RNAfold server. It shows the scale of occurrence for each nucleotide interaction.

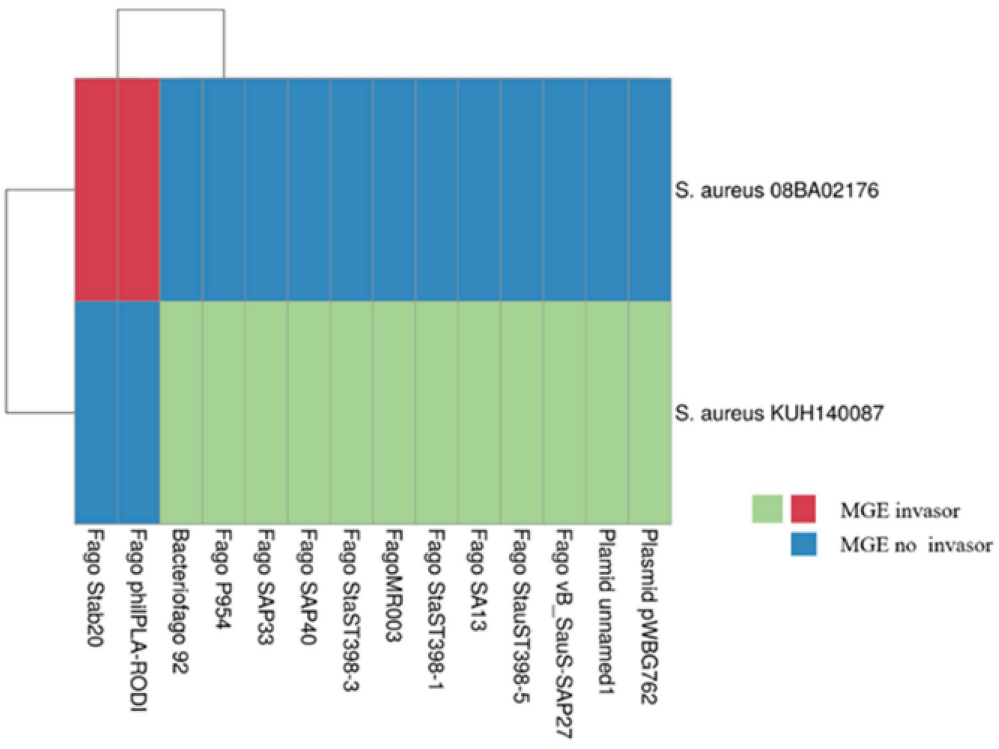


FIGURE 5 | CRISPR-Cas ported strains that exhibited information of known mobile genetic elements (MGE). The annotations of the heat map were appended. The strains shown are unique, in which some of the spacer sequences were associated with an MGE through BLAST. The map was obtained in ClusVis.

bacteriophages (Naser et al., 2017). The existence of the CRISPR-Cas system in a minimalist form, inactive, partially active, or active in MGE is the result of the constant coevolution between microorganisms and MGE (Faure et al., 2019), or due to competency between plasmids as a plasmid incompatibility mechanism (Kamruzzaman and Iredell, 2020).

In this study, the CRISPR-Cas system was found in six *S. aureus* strains. Interestingly, the strains were isolated from different countries: *S. aureus* 08BA02176 in Canada (Golding et al., 2012), *S. aureus* KUH140087 in Kyoto, Japan (Hikichi et al., 2019), *S. aureus* JS395 in Switzerland (Larsen et al., 2017), and the *S. aureus* AR_0472, AR_0470, and AR_0473 strains, whose geographical origin is unknown. The few CRISPR-Cas-bearing strains and their different geographical origin led us to think that the CRISPR-Cas system in *S. aureus* might be a spontaneous biological phenomenon, which means that the CRISPR-Cas system found in this study might be part of a bacterium that lives together with *S. aureus*. It has been demonstrated that in *S. aureus* 08BA02176, *S. schleiferi* TSCC54, and *S. capitis* CR01, the CRISPR-Cas system is inside the staphylococcal chromosomic cassette (*SSC*) *SSCmec*. The *SSCmec* is flanked by insertion sequences (IS), in *S. aureus* 08BA02176 by IS6 and ISL3, in *S. schleiferi* TSCC54 by IS6 and IS1182, and in *S. capitis* CR01 by an IS6, and the presence of the MGE mentioned indicates that the CRISPR-Cas system has been transferred horizontally to other strains and species of *Staphylococcus* (Rossi et al., 2017). The results of this study support the proposal of Rossi et al. (2017) and allow us to postulate that the CRISPR-Cas in *S. aureus* might be a spontaneous event consequence of a horizontal transfer of the *SCCmec* because of the low number of strains harboring the CRISPR-Cas system and their different geographical regions. Further evidence of horizontal transfer of the CRISPR-Cas system through *SCCmec* requires additional bioinformatic analysis and its *in vitro* demonstration.

The CRISPR-Cas systems in the *S. aureus* strains analyzed in this study are classified as subtype III-A, since the gen *cas10/csm1* is found (Koonin et al., 2017). Studies have been demonstrated that the HD dominion of the protein Cas10/Csm1 is responsible for the activity ssDNasa and the protein Csm3 of the activity endoribonucleases (Tamulaitis et al., 2017). The crRNA is essential for the operation of the CRISPR-Cas system (Behler and Hess, 2020). **Figures 3A,B** show that the Cas protein and the SRc coevolution comply with the correct functioning of the CRISPR-Cas system and that the *stem-loop* structure is conserved; in the alignment (**Figure 2B**) of the SRc, it has been demonstrated that the presence of conserved motifs is formed by four cysteines and four guanines that flanked an inner region of eight nucleotides. Motifs rich C and G can interact to form a pair of C:G, which has also been observed in *Proteus* spp. (Qu et al., 2019). The alignments of SRc evidence the conserved motifs (Yang et al., 2020) that interact to generate secondary structures (**Figure 4**), which are characterized by the *stem-loop* structure (Bhaya et al., 2011) that serves as a point to process the pre-crRNA through the endonuclease Cas6 (Wakefield et al., 2015). The secondary structure stability is bigger as far as there are more G:C interactions and less MFE (Trotta, 2014); nevertheless, the nucleotides bound in different forms to G:C, so there are also stable structures (Yang et al., 2015; Negahdaripour et al., 2017).

Multidrug-resistant (MDR) strains arise because exposition to antimicrobial compound (AMC) in the environment selects them (Vuotto et al., 2018; Sanderson et al., 2020), as well as horizontal AMC gene transfer (Zarei-Baygi et al., 2019) through MGE (Baker et al., 2018). This relation between ARG and MGE is difficult for the therapy of MDR bacterial infections (Vuotto et al., 2018). *S. aureus* strains that contain the CRISPR-Cas system are detected in this study, three are from a clinical origin (08BA02176, JS395, and KUH140087), and the origin of the rest (AR_0470, AR_0472, and AR_0473) is unknown. The presence of antimicrobial resistance (AMR) in the environment may be the result of its incorrect use, for instance, the livestock industry and the pig industry, where they are used for animal breeding (Zhu et al., 2013) as well as their indiscriminate use to treat infection diseases (Saha et al., 2019) or their long-lasting use in severe or chronic treatments (Karaiskos et al., 2019). The cross pollution favors the outcome of MDR to different places far from its origin (Uhlemann et al., 2017; Aeksiri et al., 2019; Cohen et al., 2019). The effort and the economic consumption to the development of antimicrobial products (Chung and Khanum, 2017; Hashemi et al., 2018), mainly those are effective against MDR strains with metal in the form of a nanoparticle (Alavi and Rai, 2019; Heidary et al., 2019; Kumar et al., 2019), have promoted the search of new treatments, particularly the treatment of bacteriophages (Wernicki et al., 2017). The bacteriophages are being considered as an alternative to therapy in *S. aureus* MDR strains (MRSA), *S. haemolyticus* (MRSH), and *Staphylococcus epidermidis* (MRSE) infections (Oduor et al., 2020). However, in this study, we found that the CRISPR-Cas system may be a factor that could compromise the efficacy of bacteriophage therapy. The BLAST analysis of SE6 has shown that *S. aureus* 08BA02176 is capable of counteracting the Stab20 bacteriophages infection. Oduor et al. (2019) isolated the Stab20, Stab21, Stab22, and Stab23 bacteriophages, and later it was determined that Stab20 and Stab21 infected 41 and 40, among them, 100 *Staphylococcus* spp. (MRSA, MSSA, *Staphylococcus intermedius*, *S. epidermidis*, *Staphylococcus saprophyticus*, and *S. haemolyticus*) strains; moreover, it was found that Stab20 and Stab21 are better spread in some *S. aureus* strains. The Stab21 bacteriophage is capable to infect an isolated *S. aureus* from a patient with chronic sinusitis (Oduor et al., 2020). The presence of one SE that matches with Stab20 in the loci CRISPR of the 08BA02176 strain implies that infection with this strain would be difficult, or impossible to treat with a Stab20 bacteriophage therapy. Likewise, *S. aureus* 08BA02176 strain demonstrated its capability to destroy the ϕIPLA-RODI phage; this phage, when used against *S. aureus* forming a biofilm, caused a reduction of the population of *S. aureus* after 18 h (González et al., 2017); nevertheless, the presence of *S. aureus* 08BA02176 as part of the biofilm makes the use of the ϕIPLA-RODI phage difficult as a treatment. In contrast, it was demonstrated that the ϕMR003 phage infected 97% of the MRSA strains in the study of Peng et al. (2019); however, the CRISPR-Cas system of *S. aureus* KUH140087 prevents attack by the ϕMR003 phage.

Despite the ongoing protocols using the bacteriophages to treat infections caused by *S. aureus* (Kaźmierczak et al., 2014; Cui et al., 2017) in an animal model and human studies,

it is necessary to generate more knowledge about the CRISPR-Cas system in more *S. aureus* strains to develop reliable bacteriophage therapies. Nowadays, only 12 *S. aureus* strains contain the reliable CRISPR-Cas system: AH1, AH2, AH3, SH1, SH2, and SH3 strains from isolated clinics (Cao et al., 2016), and the 08BA02176, KUH140087, JS395, AR_0470, AR_0472, and AR_0473 strains found in this investigation, as well as the ones previously found in the study by Cao et al. (2016) are 08BA02176 and JS395 strains as the CRISPR-Cas system carrier.

In conclusion, we determined that the CRISPR-Cas system found has an origin from other bacteria before getting into the different *S. aureus* strains detected in this study, due to its rare presence in clinical infections and its wide geographical countries where the CRISPR-Cas system was detected; moreover, the CRISPR-Cas system-bearing bacteria can destroy the bacteriophages becoming the limiting factor that could avoid the therapeutic use of the bacteriophages. Our results can be complemented with the CRISPR-Cas system detection in more *S. aureus* strains; thus, the enrichment of the database is to associate the memory of the CRISPR-Cas system with the bacteriophages and to discriminate among the best candidates for the curative therapies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

EC-L and VB-G: design of the work. EC-L, MC-H, AM-V, and GC-E: acquisition and analysis of the data. EC-L, GR, WC-P: writing and revision of the content. EC-L, RF-M, and KV: writing of the content and contribution to figures. VB-G: approval of the last version. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.656996/full#supplementary-material

REFERENCES

Abby, S. S., Néron, B., Ménager, H., Touchon, M., and Rocha, E. P. (2014). MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS One* 9:e110726. doi: 10.1371/journal.pone.0110726

Aeksiri, N., Toanan, W., Sawikan, S., Suwannarit, R., Pungpomin, P., Khieokhajonkhet, A., et al. (2019). First detection and genomic insight into mcr-1 encoding plasmid-mediated colistin-resistance gene in *Escherichia coli* ST101 isolated from the migratory bird species *Hirundo rustica* in Thailand. *Microb. Drug Resist.* 25, 1437–1442. doi: 10.1089/mdr.2019.0020

Alavi, M., and Rai, M. (2019). Recent advances in antibacterial applications of metal nanoparticles (MNPs) and metal nanocomposites (MNCs) against multidrug-resistant (MDR) bacteria. *Expert Rev. Anti-Infect. Ther.* 17, 419–428. doi: 10.1080/14787210.2019.1614914

Baker, S., Thomson, N., Weill, F. X., and Holt, K. E. (2018). Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. *Science* 360, 733–738. doi: 10.1126/science.aar3777

Behler, J., and Hess, W. R. (2020). Approaches to study CRISPR RNA biogenesis and the key players involved. *Methods* 172, 12–26. doi: 10.1016/j.ymeth.2019.07.015

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.* 45, 273–297. doi: 10.1146/annurev-genet-110410-132430

Box, A. M., McGuffie, M. J., O'Hara, B. J., and Seed, K. D. (2015). Functional analysis of bacteriophage immunity through a type I-E CRISPR-Cas system in *Vibrio cholerae* and its application in bacteriophage genome engineering. *J. Bacteriol.* 198, 578–590. doi: 10.1128/JB.00747-15

Cao, L., Gao, C. H., Zhu, J., Zhao, L., Wu, Q., Li, M., et al. (2016). Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 306, 686–696. doi: 10.1016/j.ijmm.2016.08.005

Carpenter, M. R., Kalburge, S. S., Borowski, J. D., Peters, M. C., Colwell, R. R., and Boyd, E. F. (2017). CRISPR-Cas and contact-dependent secretion systems present on excisable pathogenicity islands with conserved recombination modules. *J. Bacteriol.* 199, e00842–e00816. doi: 10.1128/JB.00842-16

Chakraborty, S., Waise, T. M., Hassan, F., Kabir, Y., Smith, M. A., and Arif, M. (2009). Assessment of the evolutionary origin and possibility of CRISPR-Cas (CASS) mediated RNA interference pathway in *Vibrio cholerae* O395. *In Silico Biol.* 9, 245–254. doi: 10.3233/ISB-2009-0402

Chambers, H. F., and Deleo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* 7, 629–641. doi: 10.1038/nrmicro2200

Chung, P. Y., and Khanum, R. (2017). Antimicrobial peptides as potential anti-biofilm agents against multidrug-resistant bacteria. *J. Microbiol. Immunol. Infect.* 50, 405–410. doi: 10.1016/j.jmii.2016.12.005

Cohen, K. A., Manson, A. L., Abeel, T., Desjardins, C. A., Chapman, S. B., Hoffner, S., et al. (2019). Extensive global movement of multidrug-resistant *M. tuberculosis* strains revealed by whole-genome analysis. *Thorax* 74, 882–889. doi: 10.1136/thoraxjnl-2018-211616

Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., et al. (2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46, W246–W251. doi: 10.1093/nar/gky425

Craft, K. M., Nguyen, J. M., Berg, L. J., and Townsend, S. D. (2019). Methicillin-resistant *Staphylococcus aureus* (MRSA): antibiotic-resistance and the biofilm phenotype. *Med. Chem. Commun.* 10, 1231–1241. doi: 10.1039/C9MD00044E

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Cui, Z., Guo, X., Dong, K., Zhang, Y., Li, Q., Zhu, Y., et al. (2017). Safety assessment of *Staphylococcus* phages of the family Myoviridae based on complete genome sequences. *Sci. Rep.* 7:41259. doi: 10.1038/srep41259

Faruque, S. M., Naser, I. B., Islam, M. J., Faruque, A. S., Ghosh, A. N., Nair, G. B., et al. (2005). Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proc. Natl. Acad. Sci. U. S. A.* 102, 1702–1707. doi: 10.1073/pnas.0408992102

Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019). CRISPR-Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* 17, 513–525. doi: 10.1038/s41579-019-0204-7

Golding, G. R., Bryden, L., Levett, P. N., McDonald, R. R., Wong, A., Graham, M. R., et al. (2012). Whole-genome sequence of livestock-associated st398 methicillin-resistant *Staphylococcus aureus* isolated from humans in Canada. *J. Bacteriol.* 194, 6627–6628. doi: 10.1128/JB.01680-12

González, S., Fernández, L., Campelo, A. B., Gutiérrez, D., Martínez, B., Rodríguez, A., et al. (2017). The behavior of *Staphylococcus aureus* dual-

species biofilms treated with bacteriophage phiIPLA-RODI depends on the accompanying microorganism. *Appl. Environ. Microbiol.* 83, e02821–e02816. doi: 10.1128/AEM.02821-16

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360

Hashemi, M. M., Mmuoegbulam, A. O., Holden, B. S., Coburn, J., Wilson, J., Taylor, M. F., et al. (2018). Susceptibility of multidrug-resistant bacteria, isolated from water and plants in Nigeria, to Ceragenins. *Int. J. Environ. Res. Public Health* 15:2758. doi: 10.3390/ijerph15122758

Heidary, M., Zaker Bostanabad, S., Amini, S. M., Jafari, A., Ghalami Nobar, M., Ghodousi, A., et al. (2019). The anti-mycobacterial activity of Ag, ZnO, And Ag-ZnO nanoparticles against MDR- and XDR-*Mycobacterium tuberculosis*. *Infect. Drug Resist.* 12, 3425–3435. doi: 10.2147/IDR.S221408

Hikichi, M., Nagao, M., Murase, K., Aikawa, C., Nozawa, T., Yoshida, A., et al. (2019). Complete genome sequences of eight methicillin-resistant *Staphylococcus aureus* strains isolated from patients in Japan. *Microbiol. Resour. Announc.* 8, e01212–e01219. doi: 10.1128/MRA.01212-19

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278. doi: 10.1016/j.cell.2014.05.010

Huff, F., Muth, C., Naumer, C., and Meinhardt, F. (2017). The restriction modification system of *Bacillus licheniformis* MS1 and generation of a readily transformable deletion mutant. *Appl. Microbiol. Biotechnol.* 101, 7933–7944. doi: 10.1007/s00253-017-8532-x

Jiang, F., and Doudna, J. A. (2015). The structural biology of CRISPR-Cas systems. *Curr. Opin. Struct. Biol.* 30, 100–111. doi: 10.1016/j.sbi.2015.02.002

Jiang, F., and Doudna, J. A. (2017). CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529. doi: 10.1146/annurev-biophys-062215-010822

Kamruzzaman, M., and Iredell, J. R. (2020). CRISPR-Cas system in antibiotic resistance plasmids in *Klebsiella pneumoniae*. *Front. Microbiol.* 10:2934. doi: 10.3389/fmicb.2019.02934

Karaiskos, I., Galani, L., Sakka, V., Gkoufa, A., Sopilidis, O., Chalikopoulos, D., et al. (2019). Oral fosfomycin for the treatment of chronic bacterial prostatitis. *J. Antimicrob. Chemother.* 74, 1430–1437. doi: 10.1093/jac/dkz015

Kaźmierczak, Z., Górski, A., and Dąbrowska, K. (2014). Facing antibiotic resistance: *Staphylococcus aureus* phages as a medical tool. *Viruses* 6, 2551–2570. doi: 10.3390/v6072551

Koonin, E. V., and Makarova, K. S. (2017). Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back. *Genome Biol. Evol.* 9, 2812–2825. doi: 10.1093/gbe/evx192

Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008

Kumar, S., Majhi, R. K., Singh, A., Mishra, M., Tiwari, A., Chawla, S., et al. (2019). Carbohydrate-coated gold-silver nanoparticles for efficient elimination of multidrug resistant bacteria and in vivo wound healing. *ACS Appl. Mater. Interfaces* 11, 42998–43017. doi: 10.1021/acsami.9b17086

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Labrie, S. J., Samson, J. E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8, 317–327. doi: 10.1038/nrmicro2315

Labruère, R., Sona, A. J., and Turos, E. (2019). Anti-methicillin-resistant *Staphylococcus aureus* nanoantibiotics. *Front. Pharmacol.* 10:1121. doi: 10.3389/fphar.2019.01121

Larsen, J., Andersen, P. S., Winstel, V., and Peschel, A. (2017). *Staphylococcus aureus* CC395 harbours a novel composite staphylococcal cassette chromosome mec element. *J. Antimicrob. Chemother.* 72, 1002–1005. doi: 10.1093/jac/dkw544

Lorenz, R., Bernhart, S. H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6:26. doi: 10.1186/1748-7188-6-26

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2013). Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* 41, 4360–4377. doi: 10.1093/nar/gkt157

McGinn, J., and Marraffini, L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.* 17, 7–12. doi: 10.1038/s41579-018-0071-7

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res.* 43, W566–W570. doi: 10.1093/nar/gkv468

Naser, I. B., Hoque, M. M., Nahid, M. A., Tareq, T. M., Rocky, M. K., and Faruque, S. M. (2017). Analysis of the CRISPR-Cas system in bacteriophages active on epidemic strains of *Vibrio cholerae* in Bangladesh. *Sci. Rep.* 7:14880. doi: 10.1038/s41598-017-14839-2

Negahdaripour, M., Nezafat, N., Hajighahramani, N., Rahmatabadi, S. S., and Ghasemi, Y. (2017). Investigating CRISPR-Cas systems in *Clostridium botulinum* via bioinformatics tools. *Infect. Genet. Evol.* 54, 355–373. doi: 10.1016/j.meegid.2017.06.027

Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820

Oduor, J., Kadija, E., Nyachieo, A., Mureithi, M. W., and Skurnik, M. (2020). Bioprospecting *Staphylococcus* phages with therapeutic and bio-control potential. *Viruses* 12:133. doi: 10.3390/v12020133

Oduor, J., Kiljunen, S., Kadija, E., Mureithi, M. W., Nyachieo, A., and Skurnik, M. (2019). Genomic characterization of four novel *Staphylococcus myoviruses*. *Arch. Virol.* 164, 2171–2173. doi: 10.1007/s00705-019-04267-0

Ostria-Hernández, M. L., Sánchez-Vallejo, C. J., Ibarra-García, J. A., and Castro-Escarpulli, G. (2015). Survey of clustered regularly interspaced short palindromic repeats and their associated Cas proteins (CRISPR/Cas) systems in multiple sequenced strains of *Klebsiella pneumoniae*. *BMC Res. Notes* 8:332. doi: 10.1186/s13104-015-1285-7

Peng, C., Hanawa, T., Azam, A. H., LeBlanc, C., Ung, P., Matsuda, T., et al. (2019). Silviavirus phage ϕMR003 displays a broad host range against methicillin-resistant *Staphylococcus aureus* of human origin. *Appl. Microbiol. Biotechnol.* 103, 7751–7765. doi: 10.1007/s00253-019-10039-2

Pourcel, C., Touchon, M., Villeriot, N., Vernadet, J. P., Couvin, D., Toffano-Nioche, C., et al. (2020). CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and *cas* genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res.* 48, D535–D544. doi: 10.1093/nar/gkz915

Qu, D., Lu, S., Wang, P., Jiang, M., Yi, S., and Han, J. (2019). Analysis of CRISPR/Cas system of *Proteus* and the factors affected the functional mechanism. *Life Sci.* 231:116531. doi: 10.1016/j.lfs.2019.06.006

Rose, R., Golosova, O., Sukhomlinov, D., Tiunov, A., and Prosperi, M. (2019). Flexible design of multiple metagenomics classification pipelines with UGENE. *Bioinformatics* 35, 1963–1965. doi: 10.1093/bioinformatics/bty901

Rossi, C. C., Andrade-Oliveira, A. L., and Giambiagi-deMarval, M. (2019). CRISPR tracking reveals global spreading of antimicrobial resistance genes by *Staphylococcus* of canine origin. *Vet. Microbiol.* 232, 65–69. doi: 10.1016/j.vetmic.2019.04.009

Rossi, C. C., Souza-Silva, T., Araújo-Alves, A. V., and Giambiagi-deMarval, M. (2017). CRISPR-Cas systems features and the gene-reservoir role of coagulase-negative staphylococci. *Front. Microbiol.* 8:1545. doi: 10.3389/fmicb.2017.01545

Saha, S., Tariq, R., Tosh, P. K., Pardi, D. S., and Khanna, S. (2019). Faecal microbiota transplantation for eradicating carriage of multidrug-resistant organisms: a systematic review. *Clin. Microbiol. Infect.* 25, 958–963. doi: 10.1016/j.cmi.2019.04.006

Sanderson, H., Ortega-Polo, R., Zaheer, R., Goji, N., Amoako, K. K., Brown, R. S., et al. (2020). Comparative genomics of multidrug-resistant Enterococcus spp. isolated from wastewater treatment plants. *BMC Microbiol.* 20:20. doi: 10.1186/s12866-019-1683-4

Schneider, T. D., and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18, 6097–6100. doi: 10.1093/nar/18.20.6097

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* 15, 169–182. doi: 10.1038/nrmicro.2016.184

Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039

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

Trotta, E. (2014). On the normalization of the minimum free energy of RNAs by sequence length. *PLoS One* 9:e113380. doi: 10.1371/journal.pone.0113380

Uhlemann, A. C., McAdam, P. R., Sullivan, S. B., Knox, J. R., Khiabanian, H., Rabadan, R., et al. (2017). Evolutionary dynamics of pandemic methicillin-sensitive *Staphylococcus aureus* ST398 and its international spread via routes of human migration. *mBio* 8, e01375–e01316. doi: 10.1128/mBio.01375-16

Vuotto, C., Donelli, G., Buckley, A., and Chilton, C. (2018). *Clostridium difficile* Biofilm. *Adv. Exp. Med. Biol.* 1050, 97–115. doi: 10.1007/978-3-319-72799-8_7

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

Wang, S., Yang, B., Ross, R. P., Stanton, C., Zhao, J., Zhang, H., et al. (2020). Comparative genomics analysis of *Lactobacillus ruminis* from different niches. *Genes* 11:70. doi: 10.3390/genes11010070

Wernicki, A., Nowaczek, A., and Urban-Chmiel, R. (2017). Bacteriophage therapy to combat bacterial infections in poultry. *Virol. J.* 14:179. doi: 10.1186/s12985-017-0849-7

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

Yang, L., Li, W., Ujiroghene, O. J., Yang, Y., Lu, J., Zhang, S., et al. (2020). Occurrence and diversity of CRISPR loci in *Lactobacillus casei* group. *Front. Microbiol.* 11:624. doi: 10.3389/fmicb.2020.00624

Yang, C., Lin, F., Li, Q., Li, T., and Zhao, J. (2015). Comparative genomics reveals diversified CRISPR-Cas systems of globally distributed *Microcystis aeruginosa*, a freshwater bloom-forming cyanobacterium. *Front. Microbiol.* 6:394. doi: 10.3389/fmicb.2015.00394

Yang, H., Xu, J., Li, W., Wang, S., Li, J., Yu, J., et al. (2018). *Staphylococcus aureus* virulence attenuation and immune clearance mediated by a phage lysin-derived protein. *EMBO J.* 37:e98045. doi: 10.15252/embj.201798045

Zarei-Baygi, A., Harb, M., Wang, P., Stadler, L. B., and Smith, A. L. (2019). Evaluating antibiotic resistance gene correlations with antibiotic exposure conditions in anaerobic membrane bioreactors. *Environ. Sci. Technol.* 53, 3599–3609. doi: 10.1021/acs.est.9b00798

Zhao, X., Yu, Z., and Xu, Z. (2018). Study the features of 57 confirmed CRISPR loci in 38 strains of *Staphylococcus aureus*. *Front. Microbiol.* 9:1591. doi: 10.3389/fmicb.2018.01591

Zhu, Y. G., Johnson, T. A., Su, J. Q., Qiao, M., Guo, G. X., Stedtfeld, R. D., et al. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3435–3440. doi: 10.1073/pnas.1222743110

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fmicb.2021.536357

Application of CRISPR/Cas9 System for Plasmid Elimination and Bacterial Killing of *Bacillus cereus* Group Strains

Xiaojing Wang[1†], *Yufei Lyu*[1†], *Siya Wang*[1,2†], *Qingfang Zheng*[1,3†], *Erling Feng*[1], *Li Zhu*[1], *Chao Pan*[1], *Shenghou Wang*[2], *Dongshu Wang*[1*], *Xiankai Liu*[1*] *and Hengliang Wang*[1*]

[1] *State Key Laboratory of Pathogens and Biosecurity, Beijing Institute of Biotechnology, Beijing, China,* [2] *Experimental Teaching Center, Shenyang Normal University, Shenyang, China,* [3] *College of Food Science and Technology, Shanghai Ocean University, Ministry of Agriculture Shanghai Engineering Research Center of Aquatic Product Processing & Preservation, Laboratory of Quality & Safety Risk Assessment for Aquatic Product on Storage and Preservation, Shanghai, China*

OPEN ACCESS

Edited by:
Carlos Alberto Moreira-Filho, University of São Paulo, Brazil

Reviewed by:
Michael Benedik, Texas A&M University, United States
Graciela Castro Escarpulli, Instituto Politécnico Nacional de México (IPN), Mexico

***Correspondence:**
Dongshu Wang
wangdongshu@foxmail.com
Xiankai Liu
liuxk007@163.com
Hengliang Wang
wanghl@bmi.ac.cn

[†] *These authors share first authorship*

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *19 February 2020*
Accepted: *19 May 2021*
Published: *10 June 2021*

Citation:
Wang X, Lyu Y, Wang S, Zheng Q, Feng E, Zhu L, Pan C, Wang S, Wang D, Liu X and Wang H (2021) Application of CRISPR/Cas9 System for Plasmid Elimination and Bacterial Killing of Bacillus cereus Group Strains. Front. Microbiol. 12:536357. doi: 10.3389/fmicb.2021.536357

The CRISPR-Cas system has been widely applied in prokaryotic genome editing with its high efficiency and easy operation. We constructed some "scissors plasmids" via using the temperature-sensitive pJOE8999 shuttle plasmid, which carry the different 20nt (N20) guiding the Cas9 nuclease as a scissors to break the target DNA. We successfully used scissors plasmids to eliminate native plasmids from *Bacillus anthracis* and *Bacillus cereus*, and specifically killed *B. anthracis.* When curing pXO1 and pXO2 virulence plasmids from *B. anthracis* A16PI2 and A16Q1, respectively, we found that the plasmid elimination percentage was slightly higher when the sgRNA targeted the replication initiation region (96–100%), rather than the non-replication initiation region (88–92%). We also tried using a mixture of two scissors plasmids to simultaneously eliminate pXO1 and pXO2 plasmids from *B. anthracis,* and the single and double plasmid-cured rates were 29 and 14%, respectively. To our surprise, when we used the scissor plasmid containing two tandem sgRNAs to cure the target plasmids pXO1 and pXO2 from wild strain *B. anthracis* A16 simultaneously, only the second sgRNA could guide Cas9 to cleave the target plasmid with high efficiency, while the first sgRNA didn't work in all the experiments we designed. When we used the CRISPR/cas9 system to eliminate the pCE1 mega-virulence plasmid from *B. cereus* BC307 by simply changing the sgRNA, we also obtained a plasmid-cured isogenic strain at a very high elimination rate (69%). The sterilization efficiency of *B. anthracis* was about 93%, which is similar to the efficiency of plasmid curing, and there was no significant difference in the efficiency of among the scissors plasmids containing single sgRNA, targeting multi-sites, or single-site targeting and the two tandem sgRNA. This simple and effective curing method, which is applicable to *B. cereus* group strains, provides a new way to study these bacteria and their virulence profiles.

Keywords: CRISPR/Cas9, sgRNA, *B. cereus* group, plasmid curing, virulence plasmid, sequence-specific antimicrobials

INTRODUCTION

Bacillus anthracis is Gram-positive aerobic bacterium. It infects humans and animals through endospores dominant, causing anthrax. Anthrax is a rapidly-spreading malignant zoonotic disease with a short incubation period and a high mortality rate (Wui et al., 2013). *Bacillus anthracis* spores has been recognized as one of the three most dangerous biological weapons in the world, with potential to cause huge economic losses and bioterrorism (Weapon, 2002). *Bacillus anthracis* contains two virulence plasmids, pXO1 (181.6kb) and pXO2 (96.2kb), which encode the anthrax toxin and capsule, respectively (Ramisse et al., 1996). The pXO1 plasmid encodes anthrax toxin proteins such as the protective antigen (PA), the lethal factor, and the edema factor. The pXO2 plasmid encodes proteins involved in capsular biosynthesis and biodegradation (Levy et al., 2012). These two plasmids are critical to the pathogenicity of *B. anthracis*, and the loss of any one plasmid results in a great reduction in virulence (Agathe et al., 2003; Ariel et al., 2003). Therefore, eliminating the two virulence mega-plasmids and further examination of the pathogenic mechanism of *B. anthracis* will be important for the prevention and control of anthrax.

We sought to construct a plasmid-cured mutant strain for studying the role of plasmids in *B. anthracis*. In previous research of this area, the removal of bacterial plasmids involved chemical agents such as acridine orange, ethidium bromide, and high temperature culture or ultraviolet irradiation (Trevors, 1986), all of which have some potential problems. The first is the poor specificity; in other words, it is possible to drive out other plasmids, while the second may cause random mutations in the host chromosome during such treatments. Therefore, our laboratory used a small, high-copy plasmid to drive out plasmids based on the plasmid incompatibility principle (Wang et al., 2011; Liu et al., 2012). This method has a better specificity than the physical and chemical methods, but requires the exact information about the sequence of the origin of replication of the plasmid. This method is also time consuming.

Bacillus cereus, a Gram-positive opportunistic pathogen (Hauge, 1950), is widespread in soil, sewage, and all types of foods. *Bacillus cereus* produces a range of virulence factors, including enterotoxins and an emetic toxin that causes diarrhea and emetic types of food poisoning (Drobniewski, 1993; Arnesen et al., 2008). The data show that there have been several cases of severe and even fatal food poisoning caused by *B. cereus*. Cereulide, a heat-stable emetic toxin produced by the non-ribosomal peptide synthetase (NRPS) gene cluster on a large *B. cereus* plasmid (Ehling-Schulz et al., 2005), is thermostable (heat stable at 121°C for 20 min) and resistant to proteolytic degradation (Shinagawa et al., 1995; Agata et al., 1996). It is responsible for the emetic type of gastrointestinal disease caused by *B. cereus*; hence, curing this plasmid to construct mutants could help researchers to investigate food poisoning outbreaks from *B. cereus*.

Bacillus cereus is very closely related to *B. anthracis*; they are both the members of the *B. cereus* group, which is a term used to describe a genetically highly homogeneous subdivision of the *Bacillus* genus (Helgason, 2000). Plasmids are vitally important for the *B. cereus* group as virulence determinants. Therefore, establishing a rapid and efficient method to eliminate virulence plasmids in *B. cereus*-group strains would help to study the relationship between the virulence plasmid and chromosome, and develop a safer, more effective vaccines and drugs to prevent and treat the diseases caused by this group of bacteria.

In 2013, the CRISPR/Cas9 system was first used for genome editing of human and mouse embryonic stem cells. The Cas9 protein contains two nuclease domains that can cleave two strands of DNA (Bikard et al., 2012). Cas9 first combines with crRNA and tracrRNA to form a complex, and then binds to DNA through the protospacer adjacent motif (PAM) sequence to form an RNA–DNA complex structure and cleave the double DNA strands (Vercoe et al., 2013). As the PAM sequence is simple in structure (5′-NGG-3′), a large number of targets can be found in almost all genes, thus the CRISPR/Cas9 system is widely used (Bikard et al., 2014). It has been successfully applied to mice, pigs, zebrafish, arabidopsis, sorghum, nematodes, yeast, *Escherichia coli* and many other animals, plants and microorganisms, and has become a genome editing tool widely used in various fields of biology and medicine (Citorik et al., 2014). pJOE8999 is a CRISPR/Cas9 single plasmid system constructed by V. Müller, who used it for *B. subtilis* genome editing to construct mutants quickly and efficiently (Altenbuchner, 2016).

In this study, the CRISPR/Cas9 system was used to cure plasmids in *B. anthracis* and *B. cereus*, and specifically kill *B. anthracis.* We found that it provides a faster and more convenient experimental method for constructing a new vaccine strain compared with older methods and may also provide a new approach toward the control of *B. anthracis.*

MATERIALS AND METHODS

Plasmids and Strains

The plasmids and strains used in this study are shown in **Table 1**.

Curing Plasmids From *B. anthracis*

Construction of "Scissors Plasmids" to Cure pXO1 and pXO2 Plasmids

We designed N20 sequences against specific sequences on pXO1 and pXO2 plasmids (GenBank accession Nos. NC_007323 and AF065404, respectively). The N20-specific target sequence (**Table 2**) in the single guide RNA (sgRNA) was designed via using sgRNAcas9_V3.0_GUI software (Xie et al., 2014). To evaluate whether this method is feasible when we did not know the precise origin of DNA replication, we designed the N20 sequences that target both the possible replication initiation region and the non-replication initiation region. The N20 sequences on pXO1 and pXO2 origins of DNA replication were named O1T and O2T, while the N20 sequences on the pXO1 and pXO2 non-origins of DNA replication were named O1NT and O2NT.

Scissors plasmids were constructed using the plus TACG connector at 5′end of O1T and O2T sequences, and an AAAC linker at 5′end of O1T and O2T reverse complementary sequences, followed by synthesizing FO1T, RO1T, FO2T, RO2T, FO1NT, RO1NT, FO2NT, and RO2NT N20 oligonucleotides

TABLE 1 | Plasmids and strains used in this study.

Plasmids and Strains	Relevant genotype and characteristics	Source
Plasmids		
pJOE8999	Contains CRISPR-Cas9 system plasmid for breaking genome; Kanamycin (30 μg/mL)	Altenbuchner, 2016
pJO1T	pJOE8999 plasmid containing sgRNA sequence target to the replication initiation region of pXO1 plasmid in *Bacillus anthracis (B. anthracis)*	This study
pJO2T	pJOE8999 plasmid containing sgRNA sequence target to the replication initiation region of pXO2 plasmid in *B. anthracis*	This study
pJO1NT	pJOE8999 plasmid containing sgRNA sequence target to the non-replication initiation region of pXO1 plasmid in *B. anthracis*	This study
pJO2NT	pJOE8999 plasmid containing sgRNA sequence target to the non-replication initiation region of pXO2 plasmid in *B. anthracis*	This study
pJN1F2T	pJOE8999 plasmid successively containing O1NT, 'gRNA and O2T, the two sgRNA sharing with a promoter P_{vanP^*}.	This study
pJF2N1T	pJOE8999 plasmid successively containing O2T, 'gRNA and O1NT, the two sgRNA sharing with a promoter P_{vanP^*}.	This study
pJN1F2W	pJOE8999 plasmid successively containing O1NT, 'gRNA-ter, P_{vanP^*} and O2T, the two sgRNA with the respective promoter P_{vanP^*}, the first sgRNA including terminator.	This study
pJF2N1W	pJOE8999 plasmid successively containing O2T,'gRNA-ter, P_{vanP^*} and O1NT, the two sgRNA with the respective promoter P_{vanP^*}, the first sgRNA including terminator.	This study
pJ16ST	pJOE8999 plasmid containing sgRNA sequence target to 16S rRNA of *B. anthracis*	This study
pJART	pJOE8999 plasmid containing sgRNA sequence target to a specific DNA fragment of the chromosome of *B. anthracis*	This study
pJHNT	pJOE8999 plasmid containing sgRNA sequence target to a specific DNA fragment of the chromosome of *B. cereus*	This study
pJA16SRT	pJOE8999 plasmid containing two sgRNAs sequence targeting on 16S rRNA (16ST) and non-replication initiation region of a specific DNA fragment of the chromosome in *B. anthracis* (ART), the two sgRNA sharing with a promoter P_{vanP^*}.	This study
pJA16SRTW	pJOE8999 plasmid containing two sgRNAs sequence targeting on 16S rRNA and non- replication initiation region of a specific DNA fragment of the chromosome in *B. anthracis*, the two sgRNA with the respective promoter P_{vanP^*}, the first sgRNA including terminator.	This study
pJp1T	pJOE8999 plasmid containing sgRNA sequence target to pCE1 plasmid of *B. cereus* BC307	This study
Strains		
B. anthracis A16PI2	pXO2 plasmid-cured derivative of wild type A16 using Plasmid Incompatibility; $pXO1^+$, $pXO2^-$	Wang et al., 2011
B. anthracis pJO1T/A16PI2	A16PI2 contains recombinant plasmid pJO1T; $pXO1^+$, $pJO1T^+$	This study
B. anthracis A16PI2D1	pXO1 plasmid-cured derivative of A16PI2 using CRISPR/Case9 system; $pXO1^-$	This study
B. anthracis A16Q1	pXO1 plasmid-cured derivative of wild type A16 using plasmid incompatibility; $pXO1^-$, $pXO2^+$	Liu et al., 2012
B. anthracis pJO2T/A16Q1	A16Q1 contains recombinant plasmid pJO2T; $pXO2^+$ $pJO2T^+$	This study
B. anthracis A16Q1D2	pXO2 plasmid-cured derivative of A16Q1 using CRISPR/Cas9 system; $pXO2^-$	This study
B. anthracis A16	Wild type A16 isolated from the carcass of a mule that died from anthrax in Hebei Province, China, in 1953; $pXO1^+$, $pXO2^+$	This lab
B. anthracis pJO1TpJO2T/A16	A16 contains recombinant plasmid pJO1T and pJO2T; $pXO1^+pXO2^+$ $pJO1T^+$ $pJO2T^+$	This study
pJN1F2T/A16	A16 contains recombinant plasmid pJN1F2T: $pXO1^+pXO2^+$ $pJN1F2T^+$	This study
pJF2N1T/A16	A16 contains recombinant plasmid pJF2N1T: $pXO1^+pXO2^+$ $pJF2N1T^+$	This study
pJN1F2W/A16	A16 contains recombinant plasmid pJN1F2TW: $pXO1^+pXO2^+$ $pJN1F2TW^+$	This study
pJF2N1W/A16	A16 contains recombinant plasmid pJF2N1TW: $pXO1^+pXO2^+$ $pJF2N1W^+$	This study
B. anthracis A16MD1	pXO1 plasmid-cured derivative of wild type A16 using CRISPR/Cas9 system; $pXO1^-$ $pXO2^+$	This study
B. anthracis A16MD2	pXO2 plasmid-cured derivative of wild type A16 using CRISPR/Cas9 system; $pXO1^+$ $pXO2^-$	This study
B. anthracis A16MDD	pXO1 and pXO2 plasmid-cured derivative of wild type A16 using CRISPR/Cas9 system; $pXO1^-$ $pXO2^-$	This study
pJ16ST/A16PI2	A16PI2 contains recombinant plasmid pJ16ST	This study
pJART/A16PI2	A16PI2 contains recombinant plasmid pJART	This study
pJA16sRT/A16PI2	A16PI2 contains recombinant plasmid pJ16sRT	This study
pJA16sRTW/A16PI2	A16PI2 contains recombinant plasmid pJ16sRTW	This study
B. cereus HN001	*B. cereus* isolated from food poisoning	This lab
pJART/HN001	HN001 contains recombinant plasmid pJART	This study
B. cereus BC307	*B. cereus* isolated from the vomit of patients	This lab
pJp1T/BC307	BC307 contains recombinant plasmid pJp1T	This study
B. cereus BC307Dp1	pCE1 plasmid-cured derivative of BC307 using CRISPR/Cas9 system; $pCE1^-$	This study

The semisynthetic promoter. The sgRNA transcribed from the semisynthetic promoter P_{vanP^*} interrupted by the lacZ α fragment (lacPOZ').

TABLE 2 | Oligonucleotide sequences and primers used in this study.

Name	Sequence	Description
O1T and PAM:	ATAACTTGTAATAGCCCTTT**AGG**	N20 sequence and PAM on pXO1 origin of DNA replication
O2T and PAM:	ACACAAAGTGATAGCCTAGA**TGG**	N20 sequence and PAM on pXO2 origin of DNA replication
FO1T	TACG ATAACTTGTAATAGCCCTTT	O1T sequence 5′ end plus TACG connector
RO1T	AAAC AAAGGGCTATTACAAGTTAT	O1T reverse complementary sequence 5′ end plus AAAC linker
FO2T	TACG ACACAAAGTGATAGCCTAGA	O2T sequence 5′ end plus TACG connector
RO2T	AAAC TCTAGGCTATCACTTTGTGT	O2T reverse complementary sequence 5′ end plus AAAC linker
O1NTand PAM	TATTCGATGAAGTCATACAC**TGG**	N20 sequence and PAM on pXO1 non-origin of DNA replication
O2NTand PAM	CTACTTATAAGAACAAACCG**AGG**	N20 sequence and PAM on pXO2 non-origin of DNA replication
FO1NT	TACG TATTCGATGAAGTCATACAC	O1NT sequence 5′ end plus TACG connector
RO1NT	AAAC GTGTATGACTTCATCGAATA	O1NT reverse complementary sequence 5′ end plus AAAC linker
FO2NT	TACG CTACTTATAAGAACAAACCG	O2NT sequence 5′ end plus TACG connector
RO2NT	AAAC CGGTTTGTTCTTATAAGTAG	O2NT reverse complementary sequence 5′ end plus AAAC linker
16ST and PAM	CGTGAGTGATGAAGGCTTTC**GGG**	N20 sequence targeting the 16SrNA region of *B. anthracis*
ART and PAM	ACACGGATGATAATAATTTG**TGG**	*B. anthracis* specific N20 sequence
Spacer-F	AACCATCACTGTACCTCCCA	Two BsaI outer primers on pJOE8999, verifying whether N20 is successfully linked
Spacer-R	GAGCGTTCTGAACAAATCCA	
pJOE8999-F	TAGTGTAGCCGTAGTTAGG	Specific sequence primers on pJOE8999 to verify the presence of pJOE8999
pJOE8999-R	AAAGGGAATGAGAATAGTG	
cya-F	AGGATTGATGTGCTGAAAGGAG	cya gene primer pair on pXO1
cya-R	TTCGTCTTTGTCGCCACTATC	
pXO1-7F	CGTACTGCTGGAATTGATGG	A specific gene primer pair on pXO1
pXO1-7R	GTCTTGGCTAACACCTGTATG	
pXO1-13F	AGAAATTGAGTTTGAATATGGTGAG	A specific gene primer pair on pXO1
pXO1-13R	AGGTTGGCTTACTGGAGATAC	
pXO1-16F	AGCACATGACATACGAAGAAC	A specific gene primer pair on pXO1
pXO1-16R	GAACATAAGAAGTCTGAATGGATAG	
pXO1-23F	AACTAAGACACAACGAATACTACG	A specific gene primer pair on pXO1
pXO1-23R	CATTATGTGGTCAAGATTATGGTTC	
pXO1-32F	TGAACATGAACTAGAGGAATTGG	A specific gene primer pair on pXO1
pXO1-32R	ATCTTCTGGAGTCGGATTAGC	
pXO1-42F	ATCTGTGCTGCTCGTATCG	A specific gene primer pair on pXO1
pXO1-42R	GGAATCCTGGAATGAATGATGG	
pXO1-51F	TTGCCTGAGGTTCCTGTTG	A specific gene primer pair on pXO1
pXO1-51R	GCTTTCTCTCCCTTTGTGTAAG	
pXO1-55F	CGAATGAAGGTTATTGGAATAGC	A specific gene primer pair on pXO1
pXO1-55R	CTGGATCTGGATTAGGTGTTAC	
pXO1-59F	GGACTCGGAACAACAATAACG	A specific gene primer pair on pXO1
pXO1-59R	CCTCTCCATTTCGGCTGAC	
pXO1-67F	AATGGGAATCAAAGTTTACAATCTG	A specific gene primer pair on pXO1
pXO1-67R	ACTGAACACCACCTACCTTATC	
pXO1-70F	CATACCATTACAGGAGCATCATC	A specific gene primer pair on pXO1
pXO1-70R	ACCAGGAATCGCAAGAACC	
pXO1-90F	AAGGAAGTAGAGGCAGAAGC	A specific gene primer pair on pXO1
pXO1-90R	TTAATGTGTTGGCGTTCAGG	
pXO1-95F	GTCTATCAGAAGTAGGTCATAACG	A specific gene primer pair on pXO1
pXO1-95R	TTCAGTAAGAGCCTCCATAGTAG	
pXO1-98F	GACTGGTATTTCTACTGGGTTTG	A specific gene primer pair on pXO1
pXO1-98R	GTCCTGCTTCTTGATGATGATG	
pXO1-116F	CCTTCGTTCTGGTGATATGTG	A specific gene primer pair on pXO1
pXO1-116R	AATAATATGTGGTGCCTCTTCTG	
pXO1-133F	ATTGTGGAGGATAGATTCTTTGG	A specific gene primer pair on pXO1

(Continued)

TABLE 2 | Continued

Name	Sequence	Description
pXO1-133R	TCTCGCTTGGCTAATTTCATC	
pXO1-142F	CGTGGACATCTGCTTGAAC	A specific gene primer pair on pXO1
pXO1-142R	GACGACCTTCCTCTTGATATTG	
capA-F	CGATGACGATGGGTGAC	capA gene primer pair on pXO2
capA-R	AGATTGAAGTACATGCGGATGG	
pXO2-007F	GCGATGGTGGAACAGGAATG	A specific gene primer pair on pXO2
pXO2-007R	TGCGTTGCTGCCGATATTG	
pXO2-016F	CGGTTTGGTATGAGTGAGGAAG	A specific gene primer pair on pXO2
pXO2-016R	ATTGGCTGTGGTGGTTGTTG	
pXO2-023F	TTGGGACAGGCGTTATAGAAAG	A specific gene primer pair on pXO2
pXO2-023R	GCAGCGAAGTCACTACATGG	
pXO2-027F	GTGGACTTCCTGTAACCGTAAG	A specific gene primer pair on pXO2
pXO2-027R	ATGTAATGGCTGCGTCACTTC	
pXO2-039F	GCTTCTCACTGGACACCTAATG	A specific gene primer pair on pXO2
pXO2-039R	CCACTCGTGCCAATGACTAC	
pXO2-060F	CGAAAGCAACAGGGATACAAAG	A specific gene primer pair on pXO2
pXO2-060R	AGATACTCTGCCCAACTTTCAC	
pXO2-084F	AGCGTTCAAATACAGTCACATC	A specific gene primer pair on pXO2
pXO2-084R	TTACCTTTGCGATTTCCTCATC	
pXO2-089F	AACTGACGGTGAATCCATGAAC	A specific gene primer pair on pXO2
pXO2-089R	ATTGCCTGACTAATCGCTAAGC	
pXO2-094F	CCTGGGCGTAAAGAAGATGG	A specific gene primer pair on pXO2
pXO2-094R	TCTCGTTGCGTGACATTATCG	
pXO2-097F	AAGCAACCCGTGGAGATTTC	A specific gene primer pair on pXO2
pXO2-097R	TGGATGTTCCGCACCTTTATAG	
pXO2-107F	TGGACGGAGAACAGGACTATG	A specific gene primer pair on pXO2
pXO2-107R	GGGCTTGCGGATACTCAGG	
pXO2-111F	ATACAAGCGAAGCATCAGTACC	A specific gene primer pair on pXO2
pXO2-111R	TCCATCGTTACAACCTCCATTC	
p1Tand PAM	AACTCCTAGTCAAGTACCAT**GGG**	N20 sequence and PAM on *B. cereus* BC307 pCE1 plasmid
Fp1T	TACG AACTCCTAGTCAAGTACCAT	p1T sequence 5′ end plus TACG connector
Rp1T	AAAC ATGGTACTTGACTAGGAGTT	p1T reverse complementary sequence 5′ end plus AAAC linker
P_{vanP^*}	GTGATTAGAGAATTGAGTAAAATGTACCTACG	The promoter from the pJOE8999 plasmid
‘gRNA	GCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTA TCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTT	The gRNA from the pJOE8999 plasmid
Ter	ACTCCATCTGGATTTGTTCAGAACGCTC GGTTGCCGCCGGGCGTTTTTTATCTAAAGC TTAGGCCCAGTCGAAAGACTG	The terminator of opp from the pJOE8999 plasmid
cesA-F	TTCGGTGTTACTGTGTCTG	A specific gene (cesA gene) primer pair on *B. cereus* BC307 pCE1
cesA-F	ATCGCATTCTCTTCCATCC	
cesB-F	AACTTCAACCACAGGACAA	A specific gene (cesB gene) primer pair on *B. cereus* BC307 pCE1
cesB-R	ACATTACTATACCGCCAACA	
cesC-F	CATGTCGGCTATCTTCCAG	A specific gene (cesC gene) primer pair on *B. cereus* BC307 pCE1
cesC-R	GCAACCAGATTCTCCACTT	
cesD-F	GTGACAAGACCATTAGACC	A specific gene (cesD gene) primer pair on *B. cereus* BC307 pCE1
cesD-R	ACCTGAGACGATTAGTAGTA	
cesH-F	TCTGTTGTGGCAATAGGT	A specific gene (cesH gene) primer pair on *B. cereus* BC307 pCE1
cesH-R	GGAATGATAACTCCTTGACA	
cesP-F	AGGTGTGGATGTGGAGAA	A specific gene (cesP gene) primer pair on *B. cereus* BC307 pCE1
cesP-R	GATTGTCGGTCAGCCTAC	
cesT-F	CAGGCGGAAGTGCTAATG	A specific gene (cesT gene) primer pair on *B. cereus* BC307 pCE1

(Continued)

TABLE 2 | Continued

Name	Sequence	Description
cesT-R	GTCCTCCTTCATAATGTATCAG	
p1-01F	AACCAAGCATACAGTCTCTT	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-01R	CGTTGACCACTATCACCAT	
p1-02F	CGTTCTTATGTAGCCGATGT	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-02R	GCTTCCTGTTATCACCACTA	
p1-03F	GGGTTTGGGTATCCGTAAT	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-03R	ATGATTGGCGAAGAAGTGT	
p1-04F	CAGCACCTATCCAATTACCA	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-04R	CATATTCAATCTCCATCCATCC	
p1-05F	CAGGAGACCAAGCACATC	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-05R	CAAGAATATACTCGCTCAGAC	
p1-06F	GGTGGAGGAACAGGAACT	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-06R	ATCGTCAGCAACTTCTACTT	
p1-07F	GAGAAGGCGATTGAAGGA	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-07R	CCAGAGTGTAATGTCTTGTT	
p1-08F	CGAATAGCAGAGCCTGATAT	A specific gene primer pair on *B. cereus* BC307 pCE1
p1-08R	GGTAATCCAGAAGTGAATGTAG	

The underlined part is the PAM sequence.

primers (**Table 2**). The two paired primers were annealed to obtain a double-stranded N20 oligonucleotide. The double-stranded N20 oligonucleotide product was inserted into pJOE8999 between the two *Bsa*I restriction sites. The detail of methods was shown S1 in **Supporting Information**. The recombinant plasmids were named pJO1T (from FO1T and RO1T primers), pJO2T (from FO2T and RO2T primers), pJO1NT (from FO1NT and RO1NT primers), and pJO2NT (from FO2NT and RO2NT primers) (**Table 1**).

To investigate curing two plasmids by using tandem sgRNAs, we also designed a new recombinant scissors plasmid that inserted two tandem sgRNAs into the bone vector pJOE8999 to cure both pXO1 and pXO2 from wild type *B. anthracis* A16 simultaneously. The sequence $N20_{first}$ – 'gRNA-$N20_{second}$ (or $N20_{first}$-'gRNA-ter-P_{vanP*}-$N20_{second}$) was inserted into the two *Bsa*I sites of pJOE8999 plasmid according to different order (**Figure 1C**). After ligation and transformation, PCR and sequencing were performed. The correctly constructed plasmids were named pJN1F2T, pJF2N1T, pJN1F2W, pJF2N1W, respectively.

Constructing Strains and Screening for Plasmid-Cured Colonies

We then transformed the demethylated scissors plasmids into *B. anthracis* by electroporation (500 Ω, 25 μF, 0.6 kV) (Shatalin and Neyfakh, 2005; Liu et al., 2012). For easy operation, we used attenuated strains of *B. anthracis* A16PI2 and A16Q1 to eliminate the pXO1 and pXO2, respectively (**Table 1**). The positive colonies selected by colony PCR with pJOE8999-F/R (**Table 2**) were designated pJO1T/A16PI2 ($pXO1^+$ $pJO1T^+$), pJO1NT/A16PI2 ($pXO1^+$ $pJO1NT^+$), pJO2T /A16Q1 ($pXO2^+$ $pJO2T^+$), and pJO2NT /A16Q1 ($pXO2^+$ $pJO2NT^+$).

Wild-type *B. anthracis*, harbors two large plasmids (pXO1 and pXO2) necessary for its complete virulence. To examine the two plasmids curing from *B. anthracis*, the demethylated plasmids pJO1T and pJO2T were equally mixed and simultaneously transformed into the virulent *B. anthracis* strain A16 ($pXO1^+$, $pXO2^+$; **Table 1**), and the positive colony recovered was designated pJO1TpJO2T /A16 ($pXO1^+$ $pXO2^+$ $pJO1T^+$ $pJO2T^+$).

These demethylated recombinant plasmids of pJN1F2T, pJF2N1T, pJN1F2W and pJF2N1W, containing two tandem sgRNAs targeting pXO1 and pXO2 respectively, were transformed into the virulent *B. anthracis* strain A16 ($pXO1^+$, $pXO2^+$; **Table 1**), the positive colonies recovered were designated pJN1F2T/A16, pJF2N1T/A16, pJN1F2W/A16, and pJF2N1W/A16 (**Table 1**).

The recombinant strains were grown in LB broth (containing 25 μg/ml kanamycin) at 30°C (220 rpm) for 3 h. The culture was added 0.4% D-mannose to induce Cas9 protein expression for 10 h. After induction of bacteria subculture (Altenbuchner, 2016), the plasmid-cured colonies were PCR-screened by diluting and plating the bacterial culture medium onto LB agar (containing 25 μg/ml kanamycin), and incubating it at 30°C overnight.

The pXO1-cured colonies were screened by colony PCR with *cya*-F/R primers, and authenticated using the other 17 specific primer pairs on pXO1 (**Table 1**). The pXO2-cured colonies were screened by colony PCR with *capA*-F/R primers, and authenticated using the other 12 specific primer pairs on pXO2 (**Table 1**).

We also used colony PCR with *cya*-F/R and *capA*-F/R primers to confirm the plasmid-cured strains from A16. We expected to obtain three types of strain from one experiment: a pXO1-cured strain, a pXO2-cured strain, and a dual plasmid-cured strain.

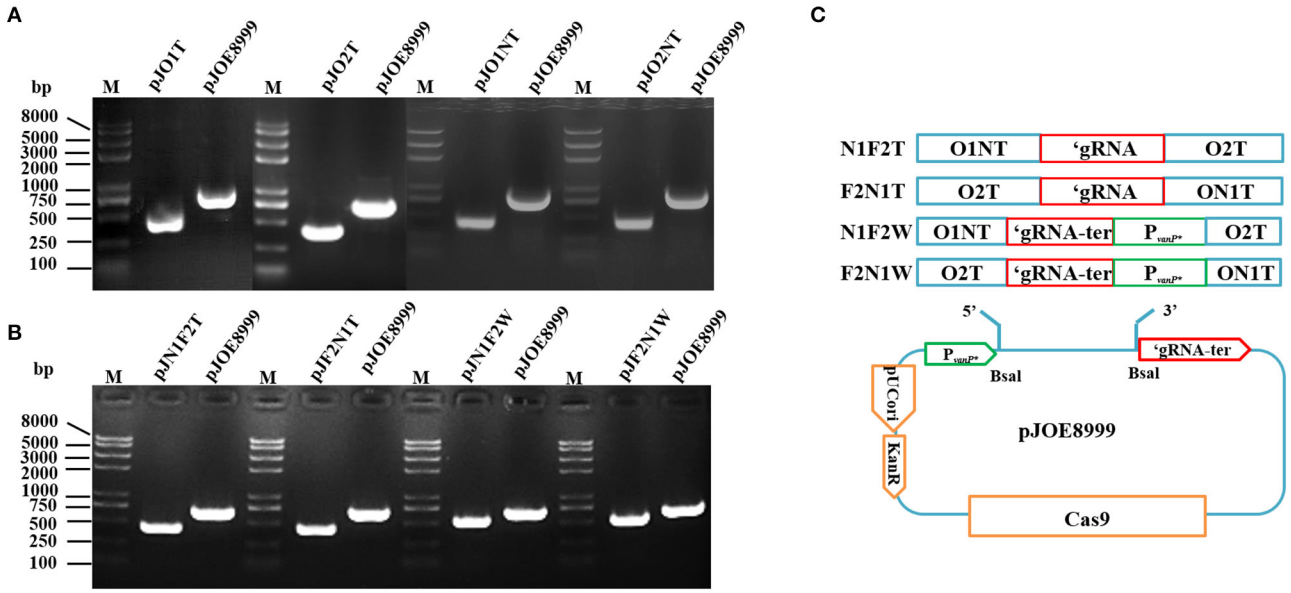


FIGURE 1 | Colony PCR for pJO1T, pJO2T, pJO1NT, pJO2NT, pJN1F2T, pJF2N1T, pJN1F2W, and pJF2N1W constructs in *Escherichia coli* DH5α with primer pair spacer-F/R. M, *Trans*2K Plus II DNA marker. **(A)** The expected sizes of the fragments from pJO1T, pJO2T, pJO1NT, and pJO2NT are smaller than those from the pJOE8999 negative control. **(B)** The expected sizes of the fragments from pJO1T, pJO2T, pJO1NT, and pJO2NT are smaller than those from the pJOE8999 negative control. **(C)** The constructing process of the recombinant plasmid with two sgRNAs of O1NT and O2T inserted into pJOE8999 in tandem.

Eliminating the Scissors Plasmids

To eliminate the scissors plasmids, we passaged the plasmid-cured colonies twice at 37°C (220 rpm) in 5 mL of LB broth separately without antibiotics. Each passage culture was diluted and spread onto agar plates without antibiotics at 30°C for 12 h. We streaked single colonies onto two agar plates with or without kanamycin, and the kanamycin-sensitive colonies were the strains that had lost the scissors plasmids. We also used colony PCR with pJOE8999-F/R to verify elimination of the scissors plasmids (**Table 2**). The pXO1 plasmid-cured strain from A16PI2 was designated A16PI2D1 (pXO1$^-$, pXO2$^-$), the pXO2 plasmid-cured strain from A16Q1 was designated A16Q1D2 (pXO1$^-$, pXO2$^-$), the pXO1 plasmid-cured strain from A16 was designated A16MD1 (pXO1$^-$, pXO2$^+$), the pXO2 plasmid-cured strain from A16 was designated A16MD2 (pXO1$^+$, pXO2$^-$), and the pXO1 and pXO2 plasmid-cured strain from A16 were designated A16MDD (pXO1$^-$, pXO2$^-$).

Western Blot Analysis of the Anthrax Toxin PA

We used western blots to verify the curing of pXO1 from *B. anthracis* A16PI2 and A16. A16PI2 (pXO1$^+$) and its derivative strain A16PI2D1 (pXO1$^-$, pXO2$^-$) comprised one group, and A16D1, A16D2, and A16DD strains comprised another group. The strains were inoculated into 50 ml of BHI broth (containing 5% horse serum and 0.8% NaHCO3) at 37°C in a 5% CO2 incubator for 13 h. The supernatant of culture was filtered through a 0.22 μm filter, and the solution was precipitated by a 3-times volume of precooled acetone for 3 h at 4°C. The acetone was evaporated, and the protein pellet dissolved in an appropriate amount of urea solution (50–200 μl) containing 8M urea, and 1% DTT. After quantification via the Bradford method, the protein samples were analyzed by SDS-PAGE and Western blots. The membrane was successively incubated with the mouse anti-PA monoclonal antibody diluted in TBST (1:25,000) for 1 h and the horseradish peroxidase (HRP)-labeled goat anti-mouse IgG (1:5,000) for 1 h. The membrane was uniformly covered with ECL luminescent liquid and photographed with a low-temperature gel imager (Kodak RP X-OMAT, USA).

Indian Ink Staining

The genes encoding capsule proteins are on the pXO2 plasmid and the pXO2-cured strain does not form the capsule. We used Indian ink staining to detect the capsule formation. A16Q1 (pXO2$^+$) and A16Q1D2 (pXO2$^-$) were one group, and A16, A16MD2 and A16MDD were the other group. The strains were inoculated onto an LB agar plate (containing 0.8% $NaHCO_3$, 5% horse serum), under 5% CO_2 at 37°C overnight. A loop of bacteria was inoculated into normal saline, a drop of Indian ink was added, and the strains were checked by phase-contrast microscopy (Eclipse TE300, Nikon, Tokyo, Japan) after squashing.

Plasmid Curing in *B. cereus*

Bacillus anthracis and *B. cereus* belong to the *B. cereus* group. This group of bacteria is mostly plasmid containing, with the bacterial virulence gene mostly located on the plasmid. We used the plasmid curing method for the same purpose as in the other members of the *B. cereus* group. This involved a clinical strain of *B. cereus* BC307(CMCC(B) 63317), which was isolated from the vomit of a patient with food poisoning. After sequencing

the whole genome by single molecule real-time sequencing using platform PacBio RS II, we found that the strain contains a 270 kb plasmid (pCE1), and that the NRPS gene cluster of the emetic cereulide toxin is located on this plasmid.

We designed the N20 sequence to target the pCE1 plasmid with the same method as above described to cure the pCE1 plasmid using the CRISPR/Cas9 system. We named this recombinant plasmid p1T. The construction method and the plasmid curing procedure were consistent with that used for *B. anthracis*. Scissors plasmid p1T was introduced into BC307 by electroporation to cure the pCE1 plasmid. After subculturing at 37°C to lose the p1T scissors plasmid, we preliminarily screened for pCE1 cured strains by colony PCR with *ces*-F/R, and further confirmed the curing of the pCE1 plasmid with 12 primer pairs (**Table 2**).

Chromosome Targeting to Specifically Kill *B. anthracis* Using the CRISPR-Cas9 System

Construction of Plasmids and Strains

We designed the N20 sequence to target the bacterial chromosome for the purpose of sterilization, which might be a new way to prevent anthrax. We designed two types of N20 sequences for specific targeting of single-site and multi-sites (16S rRNA region) on the *B. anthracis* chromosome. The constructed plasmids were confirmed by PCR and sequencing and named pJART and pJ16ST (**Table 1**). The two plasmids were transformed into *E. coli* SCS110 for demethylation. We then transformed the extracted plasmids into *B. anthracis* A16PI2 and *B. cereus* HN001 by electroporation, and designated the constructed strains pJ16ST/A16PI2, pJART/A16PI2, and pJART/HN001 (**Table 1**).

According to the logical idea of **Figure 1C**, the N20 sequences of 16S rRNA and ART were together inserted into the pJOE8999 in tandem, and the constructed plasmids were transformed into *B. anthracis* A16PI2, named the strains pJA16sRT/A16PI2 and pJA16sRTW/A16PI2.

Sterilization Efficiency Determination Using the Colony Counting Method

To evaluate the sterilization efficiency of the two plasmids (pJART and pJ16ST), the constructed strains pJ16ST/A16PI2 and pJART/A16PI2 were cultured at 28°C for 3 h, 0.4% D-mannose was added (Altenbuchner, 2016), and the culturing was continued for 12 and 24 h, respectively. Strains cultured without D-mannose were the control group. Each bacterial culture was diluted (10^{-1}-10^{-6}) and 10 μL of each one was spread onto LB agar plates to compare the number of viable bacteria.

The killing efficiency of *B. anthracis* using the recombinant plasmid with two sgRNAs of 16ST and ART inserting into pJOE8999 was assessed by the method mentioned above.

The pJART/A16PI2 and pJART/HN001 strains were cultured in LB broth (containing 25 μg/mL kanamycin) for one generation, and sub-cultured for 3 h. The second-generation cultures of the two strains were mixed and inoculated into fresh LB medium at the same concentration. D-mannose (0.4%) was added to the induced group after 3 h. The mixed cultures were diluted (10^{-6}) and spread onto LB agar plates (containing 0.5% yolk lotion, which is similar to Mannitol-Egg-Yolk-Polymyxin Agar Base) for 24 h. The specific killing efficiency of *B. anthracis* was determined as the colony forming units (CFU) for *B. anthracis* and *B. cereus*, as separately based on milky rings around the colonies on the agar plates.

Growth Curve Assays

The pJ16ST/A16PI2 and pJART/A16PI2 strains were inoculated into 5 mL LB broth (containing 25 μg/ml kanamycin) for 24 h at 37°C (220 rpm). Next, 1% of pJ16ST/A16PI2 and pJART/A16PI2 inoculum were subcultured into 35 mL of BHI broth and the optical density (OD value) was measured by the high-throughput real-time Microbial Analysis system (Gering Scientific Instruments Co. Ltd, Tianjing, China) at 37°C and 550 rpm. After 3 h, the broth of experimental group was added 0.4% D-mannose followed by monitoring for 24 h. Three independent biological repeats were performed.

Specifically Killing *B. anthracis* in *B. anthracis* and *B. cereus* Mixed Culture

The pJART/A16PI2 and pJART/HN001 strains were cultured in LB broth (containing 25 μg/mL kanamycin) for one generation, and sub-cultured for 3 h. The second-generation cultures of the two strains were mixed and inoculated into fresh LB medium at the same concentration. D-mannose (0.4%) was added to the induced group after 3 h. The mixed cultures were diluted (10^{-6}) and spread onto LB agar plates (containing 0.5% yolk lotion, which is similar to Mannitol-Egg-Yolk- Polymyxin Agar Base) for 24 h. The specific killing efficiency of *B. anthracis* was determined as the colony forming units (CFU) for *B. anthracis* and *B. cereus*, as separately based on milky rings around the colonies on the agar plates.

RESULTS

Plasmid Curing in *B. anthracis*

Identifying the Scissors Plasmids

Colony PCR was performed with spacer-F/R, and the scissors plasmid was successfully constructed. After O1T, O2T, O1NT, O2NT, N1F2T, F2N1T, N1F2W, and F2N1W were inserted into the temperature-sensitive pJOE8999 (7.8 Kb) shuttle plasmid, the amplified fragment was smaller than that of the backbone plasmid, and the DNA gel electrophoresis and DNA sequencing results showed that the "scissors plasmid" was successfully constructed (**Figure 1**). The scissors plasmids were named pJO1T, pJO2T, pJO1NT, and pJO2NT, pJN1F2T, pJF2N1T, pJN1F2W, and pJF2N1W, respectively.

Identifying the Constructed Strains

We identified the transformed *E. coli* DH5α, SCS110, and *B. anthracis* (A16PI2, A16Q1, A16) strains by colony PCR with pJOE8999-F/R primers. The result indicated that the scissors plasmid was successfully transformed into each strain (**Figure 2**). The constructed strains were named pJO1T/A16PI2 ($pXO1^+$ $pJO1T^+$), pJO2T /A16Q1

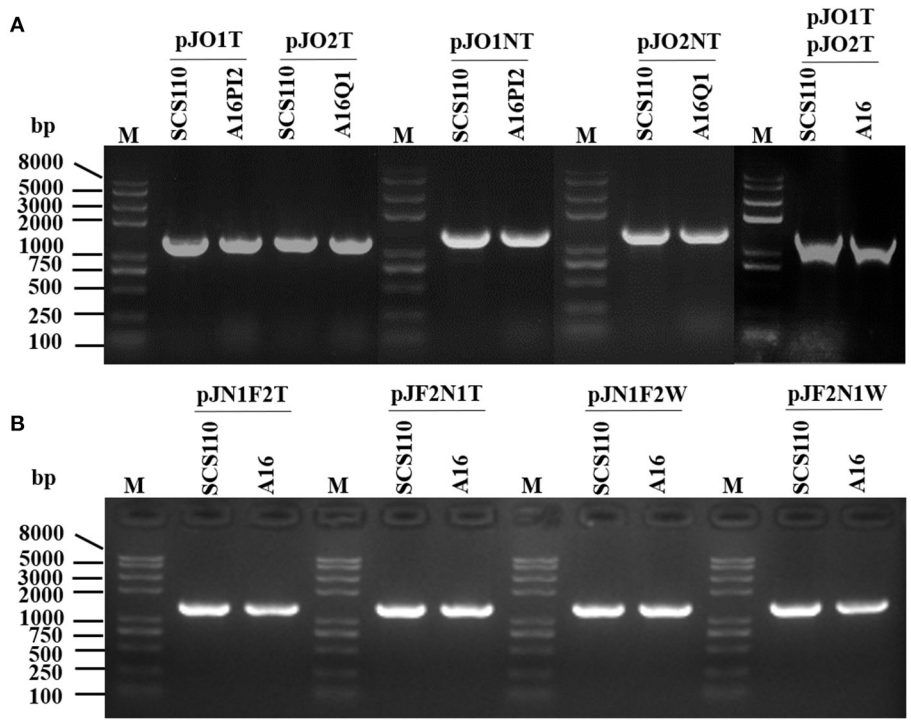


FIGURE 2 | Colony PCR for screening scissors plasmid transformants in *Bacillus anthracis* with primer pair pJOE8999-F/R. M, *Trans*2K Plus II DNA marker. **(A,B)** *Escherichia coli* SCS110 containing a recombinant plasmid was the active control. The expected sizes of the fragments from the constructed strains are consistent with the size of the amplified fragments from the active control strain.

(pXO2$^+$ pJO2T$^+$), pJO1NT/A16PI2 (pXO1$^+$ pJO1NT$^+$), pJO2NT /A16Q1 (pXO2$^+$ pJO2NT$^+$), pJO1TpJO2T /A16 (pXO1$^+$ pXO2$^+$ pJO1T $^+$ pJO2T$^+$), pJN1F2T/A16(pXO1$^+$ pXO2$^+$ pJN1F2T$^+$), pJF2N1T/A16(pXO1$^+$ pXO2$^+$ pJF2N1T$^+$), pJN1F2W/A16(pXO1$^+$ pXO2$^+$ pJN1F2W$^+$), and pJF2N1W/A16(pXO1$^+$ pXO2$^+$ pJF2N1W $^+$).

Colony PCR Screening to Identify Plasmid Curing

We used colony PCR to preliminarily screen for pXO1-cured strains from A16PI2 with *cya*-F/R primers by the scissor plasmid pJO1T or pJO1NT. The results indicated the 96% (23/24) and 92% (22/24) clones had eliminated pXO1 (**Figure 3A**). We used colony PCR to preliminarily screen for pXO2-cured strains from A16Q1 with primers *capA*-F/R by the scissor plasmids pJO2T or pJO2NT, the results of which indicated that 100% (24/24) and 88% (21/24) clones had eliminated pXO2 (**Figure 3B**). The results using the mixture plasmids of pJO1T and pJO2T to eliminate pXO1 and pXO2 from wild type strain A16 indicated that 5 clones (lanes 2–6) had eliminated the pXO1 plasmid, another 5 clones (lanes 8–12) had eliminated the pXO2 plasmid, and 2 clones (lanes 1 and 7) had eliminated both pXO1 and pXO2 (**Figure 3C**).

In the experiments of using two tandem sgRNAs to simultaneously cure the pXO1 and pXO2 in *B. anthracis* A16, the colony PCR was performed to assessed the curing efficiency with *cya*-F/R primers for the pXO1 and *capA*-F/R primers the pXO2. The 91% (10/11) and 100% (11/11) clones had eliminated pXO2 but not eliminated pXO1 by the pJN1F2T and the pJN1F2W, respectively. And the 100% (11/11) and 73% (8/11) clones had eliminated pXO1 but not eliminated pXO2 by the pJF2N1T and the pJF2N1W, respectively (**Figure 3D**). Our experimental results show that the recombinant scissors plasmid containing two tandem sgRNAs cannot simultaneously excise two target plasmids. The second sgRNA sequence could cure the corresponding target plasmid with high efficiency, but the first sgRNA didn't work in all the experiments we designed.

Elimination of Exogenous Scissors Plasmids

We used colony PCR to identify strains where the exogenous scissors plasmid had been eliminated by using pJOE8999-F/R primers (**Figure 4**). The strains that had eliminated the scissors plasmid were designated A16PI2D1, A16Q1D2, A16MD1, A16MD2, and A16MDD.

Confirmation of Plasmid Curing

Colony PCR was used this time with multiple primers to identify the cured plasmid. The A16PI2D1 pXO1-cured strain was PCR-amplified with 17 pairs of primers (collectively called pXO1-X) and the results confirmed that A16PI2D1 lacks the pXO1 plasmid (**Figure 5A**). The pXO2-cured A16Q1D2 strain was PCR-amplified with 12 pairs of primers (collectively called pXO2-X) and the results indicated that A16Q1D2 lacks the pXO2 plasmid (**Figure 5B**). A16MD1, A16MD2, and A16MDD were also PCR-amplified with 5 pairs of primers (pXO1-X) and 5

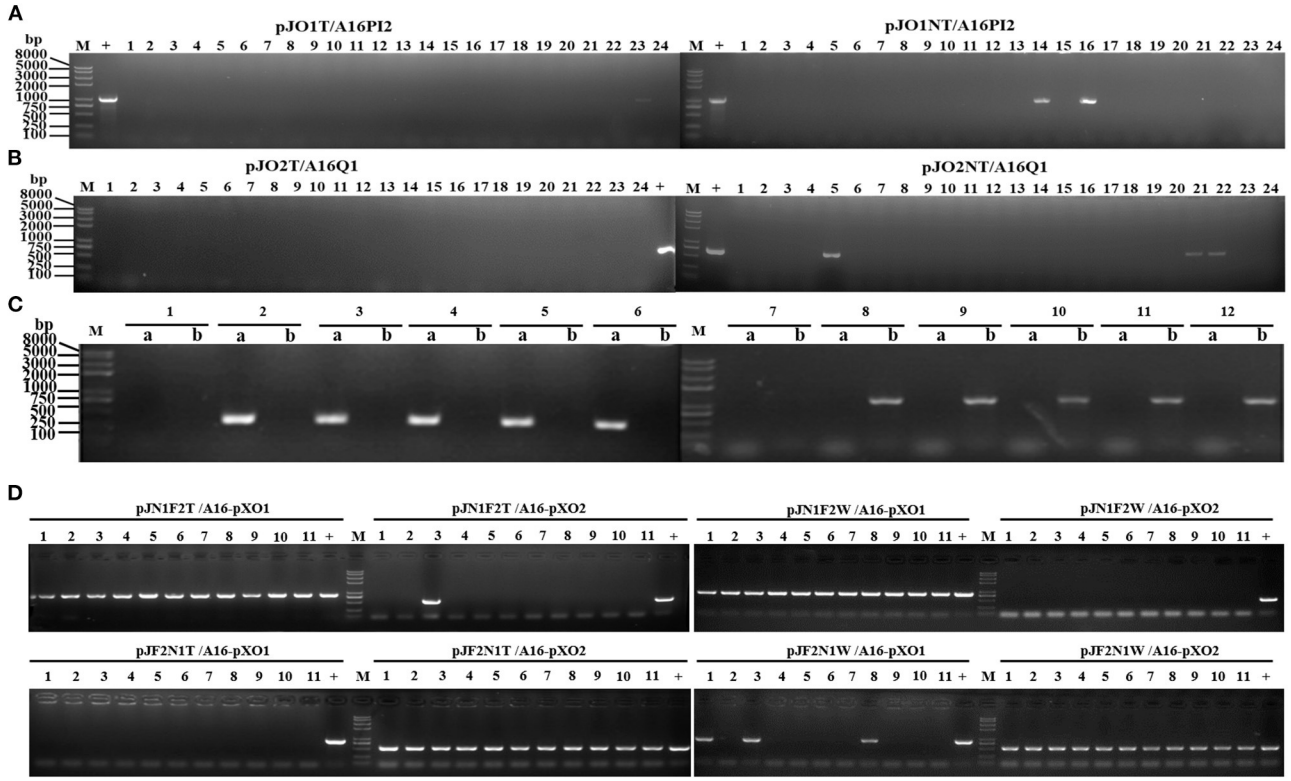


FIGURE 3 | Preliminary PCR screening of colonies for pXO1 and pXO2 plasmid-cured strains with *cya*-F/R and *capA*-F/R primers, respectively. **(A)** Colony PCR with *cya*-F/R after curing the pXO1 plasmid from pJO1T/A16PI2 and pJO1NT/A16PI2. The pXO1 plasmid-cured strains do not have a specific amplification fragment, and the efficiency of pJO1T was about 96% (expect lane 23). The efficiency of pJO1NT was about 92% (expect lanes 14, 16). **(B)** Colony PCR with capA-F/R after curing the pXO2 plasmid from pJO2T/A16Q1 and pJO2NT/A16Q1. The pXO2 plasmid-cured strains do not have a specific amplification fragment, and the efficiency of pJO2T was about 100%. The efficiency of pJO2NT was about 88% (except lanes 5, 21, 22). **(C)** Colony PCR with *capC*-F/R (lane a) and *pag*-F/R primers (lane b) to screen for plasmid-cured strains from A16. The simultaneously cured pXO1 and pXO2 plasmids in A16 do not have specific amplification fragments in lanes a and b (1 and 7 monoclonal colonies). The pXO1 plasmid-cured strains in A16 do not have a specific amplification fragment in lane b but have a specific amplification fragment in lane a (2–6 monoclonal colonies). The pXO2 plasmid-cured strains in A16 do not have a specific amplification fragment in lane a, but have a specific amplification fragment in lane b (8–12 monoclonal colonies). **(D)** The efficiency of curing plasmids from *B. anthracis* A16 by using the recombinant plasmid with two sgRNAs of O1NT and O2T inserted into pJOE8999 in tandem. Colony PCR was revealed for pXO1 and pXO2 plasmid-cured strains with *cya*-F/R and *capA*-F/R. After with 0.4% D-mannose, the random 11 clones of pJN1F2T/A16, pJN1F2W/A16, pJF2N1T/A16, and pJF2N1W/A16 were used to assaying the curing rate. The curing pXO2 efficiency of pJN1F2T/A16 and pJN1F2W/A16 were respectively ∼91 and 100%, and they both do not cure the pXO1. The curing pXO1 efficiency of pJF2N1T/A16 and pJF2N1W/A16 were respectively ∼100 and 72%, and they both do not cure the pXO2. M, *Trans*2K Plus II DNA marker; lane1-11, 11 samples; +, control.

other pairs of primers (pXO2-X) and the results indicated that A16MD1 lacks the pXO1 plasmid, A16MD2 lacks the pXO2 plasmid, and A16MDD lacks both pXO1 and pXO2 plasmids (**Figure 5C**). Thus, we successfully cured the target plasmid from the corresponding strains.

Western Blot Analysis and Indian Ink Dyeing

The pXO1 plasmid-containing strain was able to express the PA protein via the *pag* gene, whereas the pXO1 plasmid-cured one could not. Western blotting showed that A16PI2 (pXO1$^+$) and A16MD2 (pXO1$^+$ pXO2$^-$) both expressed PA (83 kD) protein, whereas A16PI2D1 (pXO1$^-$), A16MD1 (pXO1$^-$ pXO2$^+$), and A16MDD (pXO1$^-$ pXO2$^-$) did not (**Figure 6A**). Therefore, the pXO1 plasmid was cured in these strains.

The genes encoding the capsular protein are on the pXO2 plasmid. After the pXO2 plasmid was cured, the strain did not have a capsular structure. After Indian ink dyeing, the background was gray and black, and A16Q1 and A16MD1 contained the pXO2 plasmid and capsule structure, and a colorless transparent circle around the bacteria was visible against the gray background (**Figure 6B**). A16Q1D2, A16MD2, and A16MDD were cured of pXO2 plasmids, as revealed by the lack of colorless transparent circles, a well-known feature of the capsule structure (**Figure 6B**). These results indicate that plasmid pXO2 was cured in these strains.

Curing the pCE1 Plasmid in *B. cereus*

We used colony PCR with spacer-F/R primers to identify the pJp1T scissors plasmid in *E. coli* DH5α, SCS110, and *B. cereus* BC307 transformants. The constructed strain was named pJp1T/BC307 (**Figure 7A**). We then used colony PCR

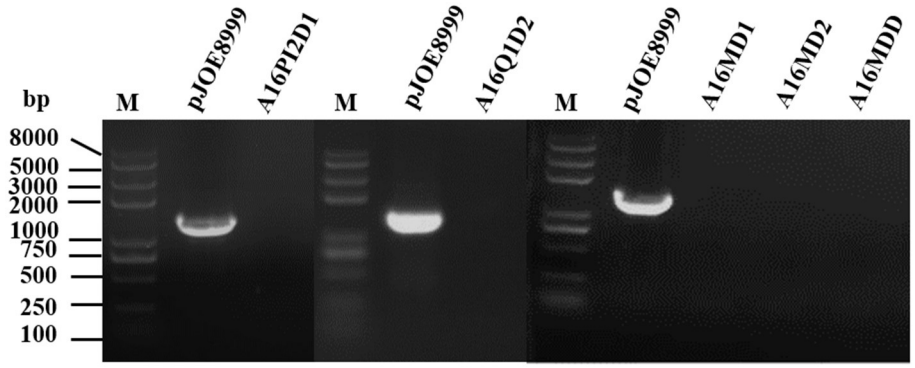


FIGURE 4 | Agarose gel electrophoretogram of colony PCR to confirm the successful elimination of the exogenous scissors plasmid. M, *Trans*2K Plus II DNA marker. The exogenous scissors plasmid-eliminated strains (A16PI2D1, A16Q1D2, A16MD1, A16MD2, and A16MDD) lack specific amplification fragments, while the active pJOE8999 control plasmid generated specific amplification fragments.

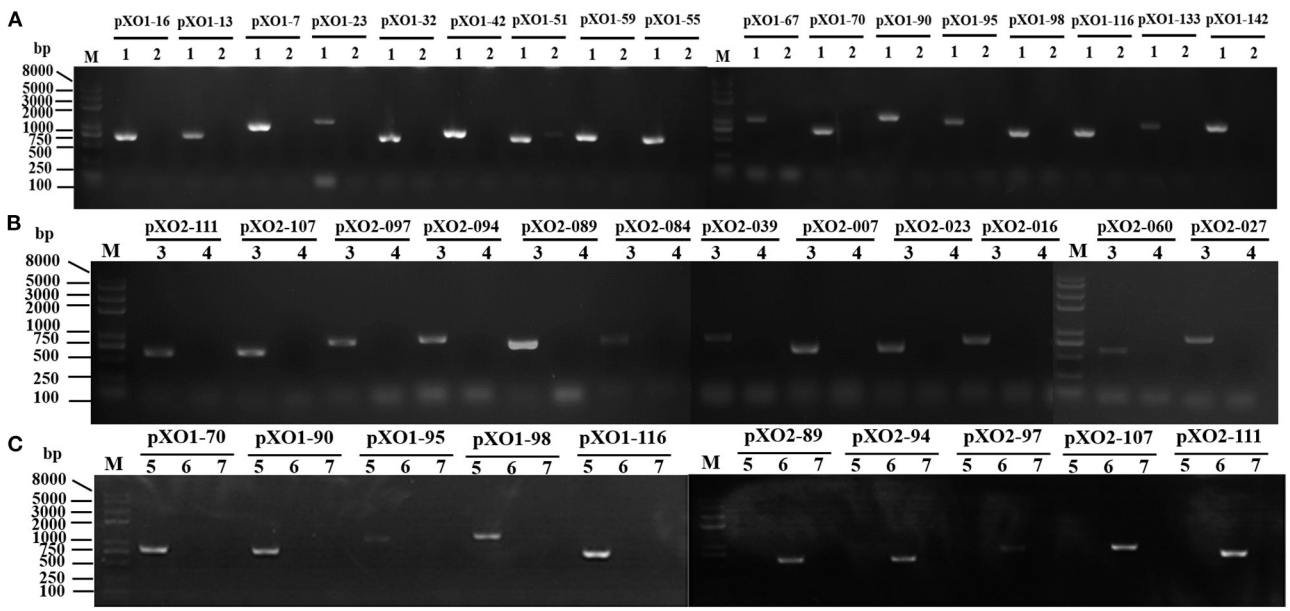


FIGURE 5 | Agarose gel electrophoresis of colony PCR to identify plasmid curing via multiple primers. **(A)** The authenticity of the A16PI2D1 pXO1-cured strain was confirmed using 17 primer pairs. "pXO1-X" is used to represent the 17 gene primers on the pXO1 plasmid. A16PI2D1 (lane 2) lacks specific amplification fragments, while the A16PI2 control (lane 1) generated specific amplification fragments. **(B)** The authenticity of the pXO2-cured A16Q1D2 strain was confirmed using 12 primer pairs. "pXO2-X" is used to represent the 12 gene primers on the pXO2 plasmid. A16Q1D2 (lane 4) lacks specific amplification fragments, while the A16Q1 control (lane 3) generated specific amplification fragments. **(C)** Plasmids cured from A16 were authenticated using pXO1-X primers and pXO2-X primers. The pXO2-cured A16MD2 ($pXO1^+pXO2^-$) strain (lane 5) lacks specific amplification fragments from pXO2-X primers, but has specific amplification fragments from pXO1-X primers. The pXO1-cured A16MD1($pXO1^-pXO2^+$) strain (lane 6) lacks specific amplification fragments from pXO1-X primers, but has specific amplification fragments from pXO2-X primers. The pXO1 and pXO2 simultaneously-cured A16MDD ($pXO1^-$ $pXO2^-$) strain (lane 7) lacks specific amplification fragments from pXO1-X primers and pXO2-X primers.

to screen for pCE1-cured strains with *ces*B-F/R primers. Monoclonal clones lacking specific amplification fragments were possibly cured the pCE1 plasmid, with an efficiency of about 68% (22/32) (**Figure 7B**). We also used colony PCR with multiple primers to determine whether the pCE1 plasmid was cured from *B. cereus* BC307. The plasmid-cured pCE1 strain was designated BC307Dp1 ($pCE1^-$) (**Figure 7C**).

Specific Killing of *B. anthracis*

Growth Curve Assays

The growth patterns of pJART/A16PI2 and pJ16ST/A16PI2 strains were continuously monitored by high-throughput real-time Microbial Analysis instrumentation. The growth of pJART/A16PI2 and pJ16ST/A16PI2 (with D-mannose) did not differ from that of the control group (without D-mannose) during the logarithmic growth phase (5–10 h), but their

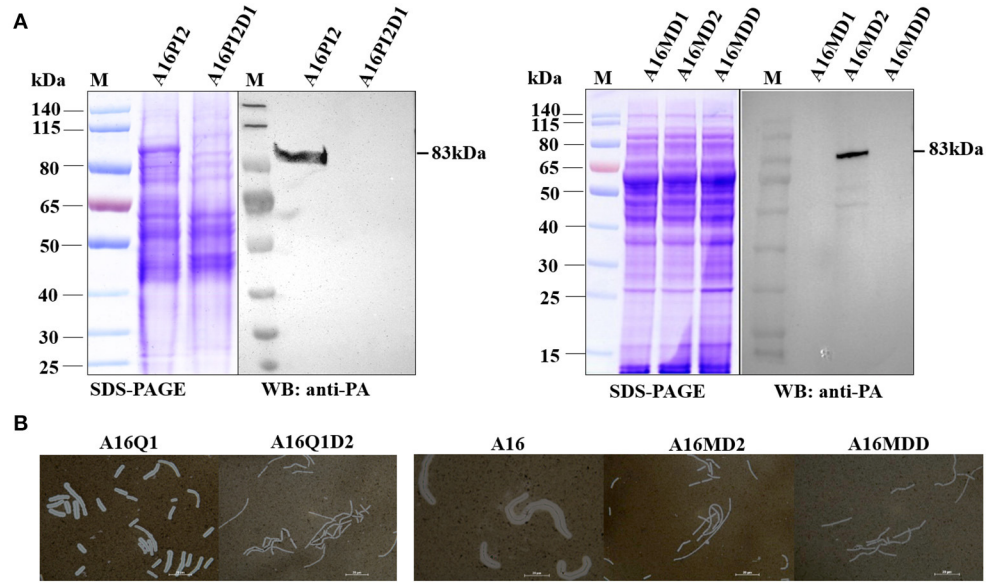


FIGURE 6 | Difference between the active strain and the plasmid-cured strain. **(A)** Western blot detection of PA expression. M, Prestained protein ladder (PageRuler Prestained Protein Ladder, Product#26616; Thermo Fisher Scientific, Vilnius, Lithuania). SDS-PAGE assay and anti-PA western blotting with two sets of samples: A16PI2 (pXO1$^+$) and A16PI2D1 (pXO1$^-$), and A16MD1 (pXO1$^-$), A16MD2 (pXO1$^+$), and A16MDD (pXO1$^-$). A16PI2D1 (pXO1$^-$) lacks the anti-PA band whereas A16PI2 (pXO1$^+$) has it. A16MD1 (pXO1$^-$) and A16MDD (pXO1$^-$) lack the anti-PA band whereas A16MD2 (pXO1$^+$) has it. **(B)** Results of bacterial capsule Indian ink staining. One set of samples were A16Q1 and A16Q1D2, and the other were A16, A16MD2, and A16MDD. After dyeing with Indian ink, the background was gray and black. A16Q1D2 lacked any colorless transparent circles around the gray-colored bacteria, whereas A16Q1 had them. A16MD2 and A16MDD lacked any colorless transparent circles around the gray-colored bacteria, whereas A16 had them. pXO2-cured A16Q1D2, A16MD2 and A16MDD lacked capsular structures.

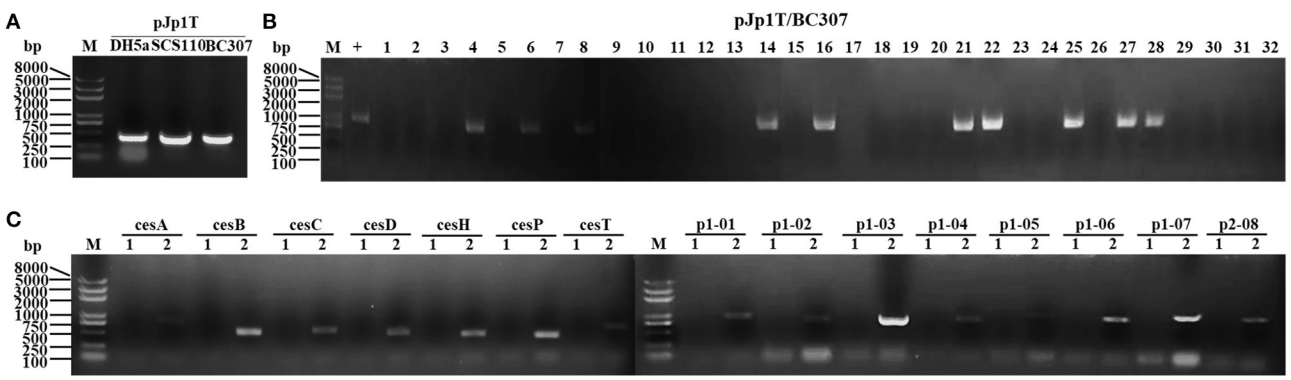


FIGURE 7 | Agarose gel electrophoresis of colony PCR for the cured pCE1 plasmid from *B. cereus*. **(A)** Colony PCR to identify the strains constructed with spacer-F/R. M, *Trans*2K Plus II DNA marker. *Escherichia coli* DH5α containing the pJp1T scissors plasmid was the control. *Bacillus cereus* BC307 containing pJp1T had the same specific amplification fragment size. **(B)** Colony PCR to screen for pCE1 plasmid-cured strains with *ces*B-F/R. M, *Trans*2K Plus II DNA marker. The pCE1 plasmid-cured strains lack specific amplification fragments, and the curing efficiency was ∼69% (lanes 1–3, 5, 7, 9–13, 15, 17–20, 23, 24, 26, and 29–32). **(C)** Colony PCR to confirm pCE1 plasmid curing using multiple primers. M, *Trans*2K Plus II DNA marker. *cesA*, *cesB*, p1-01, and p1-02 above the short line are 15 gene primers on the pCE1 plasmid. The plasmid p1-cured *B. cereus* BC307Dp1 (pCE1$^-$) strain (lane 1) lacks specific amplification fragments with *ces*X primers and p1-X primers, whereas, *B. cereus* BC307, the active control (lane 2), generated specific amplification fragments.

growth patterns decreased by OD values of 0.3–0.6 during the stationary phase (12–24 h) as compared with the control group (**Figures 8A,B**). After induction with 0.4% D-mannose, the numbers of pJART/A16PI2 colonies were (5.5 ± 2.1) × 10^5 CFU/mL and (72.5 ± 7.8) × 10^5 CFU/mL in induction group and non-induction group, respectively, with a kill efficiency of about 93% (**Figures 8C,D**). The results show that the sterilization efficiency between pJART/A16PI2 (with single-site target sgRNA) and pJ16ST/A16PI2 (with multi-site target sgRNA) was not obviously different, with both having some degree of sterilization efficiency. Under the condition of D-mannose induction, the breakage efficiency of the recombinant

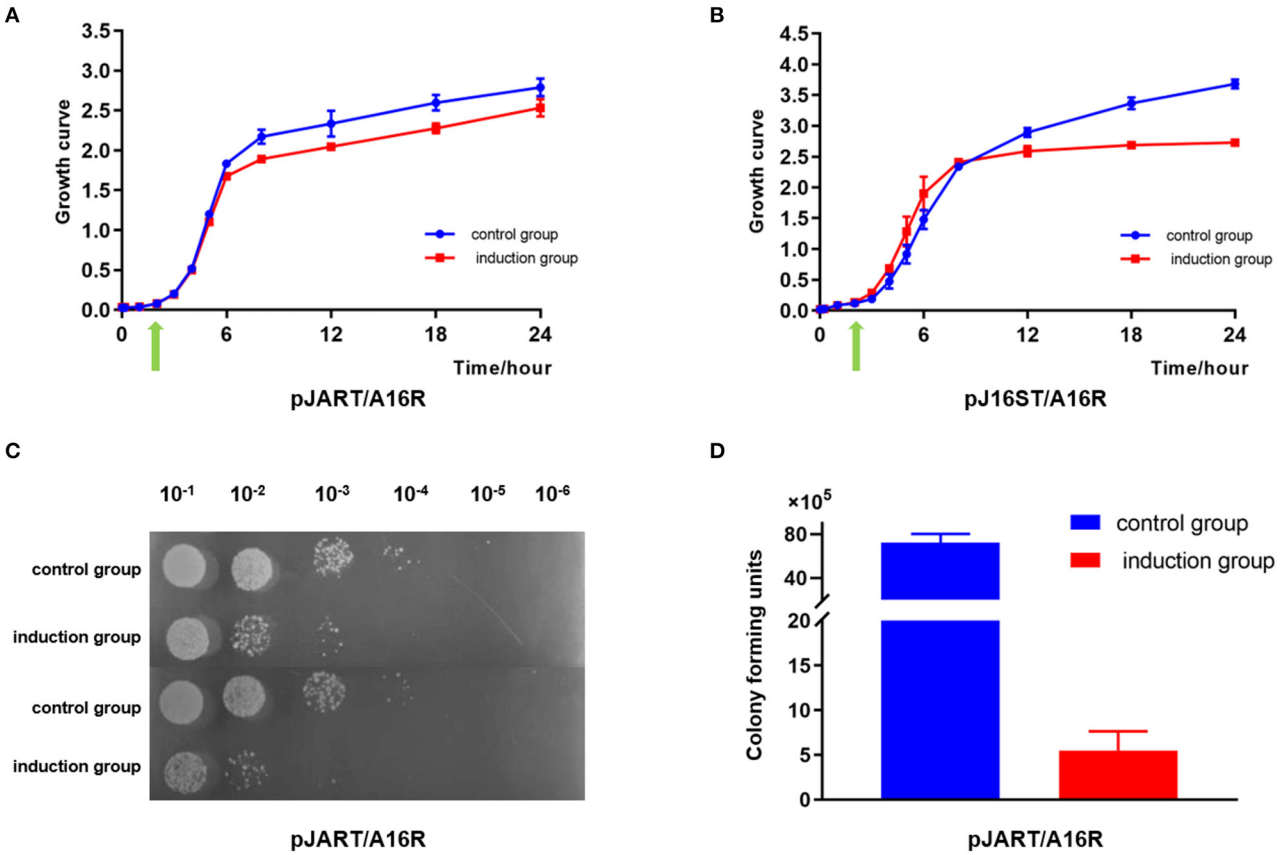


FIGURE 8 | The sterilization efficiency of *B. anthracis* by using the different sgRNAs. The experimental group was cultured continuously for 3 h and 0.4% D-mannose was added (green arrow). **(A)** pJART/A16PI2 with single-site sgRNA. The growth curve from the experimental group (pJART/A16PI2 with single-site target sgRNA, with D-mannose, red line) is lower (0.3–0.4) than that of the control group (without D-mannose, blue line) throughout the stationary phase. **(B)** pJ16ST/A16PI2 with multi-site target sgRNA. The growth curve from the experimental group (pJ16ST/A16PI2 with multi-sites sgRNA, with D-mannose, red line) is lower (0.3–0.6) than that of the control group (without D-mannose, blue line) throughout the stationary phase. **(C,D)** After incubation with 0.4% D-mannose, the numbers of pJART/A16PI2 colonies were (5.5 ± 2.1) × 10^5 CFU/mL and (72.5 ± 7.8) × 10^5 CFU/mL in the induction group and the non-induction group. Values represent the means of at least two independent replicates. Error bars represent standard deviations.

scissors plasmids pJA16sRT and pJA16sRTW, which contain two tandem sgRNAs, were the same as that of the scissor plasmids of pJ16ST and pJART, which contain only one sgRNA (see **Supplementary Figure 2**).

Specific Killing of *B. anthracis* in *B. anthracis* and *B. cereus* Mixed Cultures

Based on the above results, the specific killing efficiency of *B. anthracis* was assessed using pJART/A16PI2 and pJART/HN001 strains with single-site sgRNA. The experiment was performed according to **Figure 9A**. *Bacillus cereus* was positive for lecithinase and hemolysis activity. Before D-mannose induction, the ratio of *B. anthracis* to *B. cereus* was 57%: 43%, but after mannose induction, the ratio became 40%: 60% (**Figures 9B,C**). These results show that the pJART plasmid transfected into *B. anthracis* has the ability to specifically kill *B. anthracis* under D-mannose-induction, and the killing efficiency of *B. anthracis* was 10–17%.

DISCUSSION

The principle of "plasmid incompatibility" has been used previously to guide methodology aimed at curing the large virulence plasmid in *B. anthracis*. Plasmid incompatibility is the introduction of an incompatible plasmid group into a bacterium resulting in genetic instability of the original plasmid, possibly caused by competition for the same replication or segregation sites, or from inhibition of replication initiation (Novick and Hoppensteadt, 1978; Novick et al., 2009). Thus, knowledge of the exact origin of DNA replication in the target plasmid is essential, as is the need to culture the new plasmid-containing strain for 5–10 generations to obtain a strain that repels the large virulence plasmid. The wide application of CRISPR/Cas9 gene editing technology provides researchers with a simpler, more efficient method for plasmid curing. We only need to design different N20 sequences to guide the scissors Cas9 protein to different target sites. This method is simple to use and has a good specificity. It is not necessary to know the complete sequence or its function

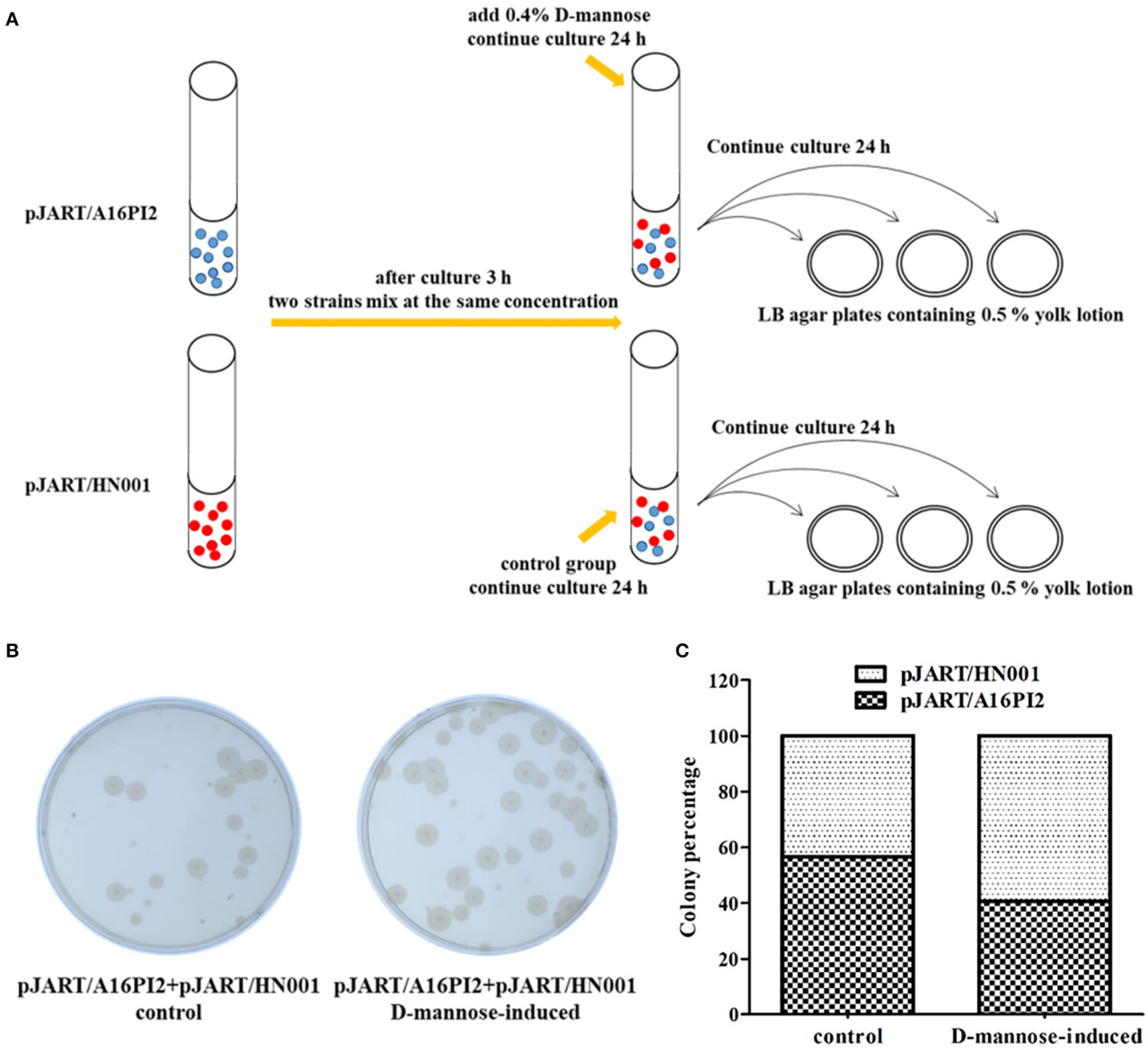


FIGURE 9 | Efficiency of *B. anthracis* sterilization in a mixed culture of *B. anthracis* and *B. cereus* under D-mannose induction. **(A)** The colony counting process used for the mixed strains. Blue balls represent strain pJART/A16PI2 and red balls represent strain pJART/HN001. **(B)** pJART/A16PI2 and pJART/HN001 grew on LB agar plates (containing 0.5% yolk lotion). The percentage of *B. anthracis* A16PI2 in the D-mannose-induced group (without proteolytic rings around the colonies) was about 10-fold less than that of the control group. **(C)** Counts of *B. anthracis* in the different groups. The black-shaded block denotes *B. anthracis* pJART/A16PI2 and the white-shaded block denotes *B. cereus* pJART/HN001. *Bacillus anthracis* (black-shaded block) accounted for 40% of the bacterial percentage in the D-mannose-induced group, while *B. anthracis* accounted for 57% in the control group.

when designing a plasmid curing protocol, and the cure efficiency is very high. In this study, the pXO1-cure efficiency was 96% when the sgRNA targeted the replication initiation region and 92% when the sgRNA targeted the non-replication initiation region. When we cured the pXO2 plasmid, the situation was much the same, with the pXO2-cure efficiency being 100% when the sgRNA targeted the replication initiation region, and 88% when the sgRNA targeted the non-replication initiation region. This indicates that there is a very slight elimination efficiency difference when the sgRNA target the replication initiation region. We also tried using a mixture of the two scissors plasmids to simultaneously eliminate both pXO1 and pXO2 virulence plasmids from *B. anthracis,* the result of which was that the single-plasmid cure rate and the double-plasmid cure rate was 29 and 14%, respectively (**Table 3**).

We also designed a new recombinant scissors plasmid that inserted two tandem sgRNAs into the bone vector pJOE8999 to cure both pXO1 and pXO2 from wild type

TABLE 3 | Plasmid elimination rates in *B. anthracis* and *B. cereus* using the CRISPR/Cas9 system.

Scissors Plasmids	Target Strain	Target Plasmid	Target ORI	Result		
				Cured	No-cured	Eliminate rate (%)
pJO1T	A16Pl2	pXO1	Yes	23	1	96
pJO1NT			No	22	24	92
pJO2T	A16Q1	PXO2	Yes	24	24	100
pJO2NT			No	21	24	88
pJO1T+pJO2T	A16	pXO1	Yes	5	12	29
		pXO2		5	12	29
		pXO1+pXO2		2	12	14
pJN1F2T	A16	pXO1	No	0	0	0
		pXO2	Yes	10	1	91
		pXO1+pXO2		0	0	0
pJF2N1T	A16	pXO1	No	11	0	100
		pXO2	Yes	0	0	0
		pXO1+pXO2		0	0	0
pJN1F2W	A16	pXO1	No	0	0	0
		pXO2	Yes	11	0	100
		pXO1+pXO2		0	0	0
pJF2N1W	A16	pXO1	No	8	3	73
		pXO2	Yes	0	0	0
		pXO1+pXO2		0	0	0
pJp1T	BC307	BC307 pCE1	No	22	10	69

B. anthracis A16 simultaneously. According to our design, the RNA sequence $N20_{pXO1}$gRNA-$N20_{pXO2}$gRNA will be obtained after transcription. This RNA sequence will be treated with RNase III or other enzymes in the bacteria to obtain two independent sgRNAs: $N20_{pXO1}$gRNA and $N20_{pXO2}$gRNA. Under the guidance of $N20_{pXO1}$ and $N20_{pXO2}$, the plasmids pXO1, and pXO2 will be targeted and cleaved respectively. However, the experimental results were inconsistent with our expectations. Our experimental results showed that the recombinant scissors plasmid containing two tandem sgRNAs could not simultaneously excise two target plasmids. The second sgRNA sequence could cure the corresponding target plasmid with high efficiency, but the first sgRNA did not work in all the experiments we designed.

We presume that the enzymes required for digestion the tandem sgRNA are not worked in *B. anthracis*, resulting in the two transcribed sgRNAs always being in tandem. In this tandem sgRNA, the 3′ end of the second sgRNA can form the structure required for binding Cas9, so the plasmid targeted by the second sgRNA can be excised. However, due to the influence of the sequence of the second sgRNA at the 3′ end of the first sgRNA, the first might not be able to form the structure required for binding to Cas9, and thus cannot cut the target plasmid. This phenomenon is interesting and worthy of further investigation.

Bacillus anthracis, *B. cereus* and *B. thuringiens*is are *B. cereus* group members, and the bacteria in this group mostly contain plasmids. Many specific biochemical functions, such as toxin production and resistance to antibacterial drugs, for example, are inherited through plasmids (Helgason, 2000). When we used the CRISPR/cas9 system to eliminate the pCE1 virulence megaplasmid from *B. cereus* BC307 by simply changing the sgRNA, we also obtained a plasmid-cured isogenic strain with a very high elimination rate (69%) (**Table 3**). We quickly and easily cured the plasmids from these strains through the CRISPR/Cas9 system, which provides new methods and ideas for studying virulence-related genes.

Although we were able to cure the plasmid by designing sgRNA, we did not know whether we could kill the bacteria by targeting 16sRNA or other sites on the chromosomes. Therefore, we designed two different target sequences with which to break the *B. anthracis* chromosome, and found that the sterilization efficiency of *B. anthracis* was about 93%, with no significant difference in the efficiency of multi-site and single-site targeting. At the same time, we also designed an experiment to sterilize the *B. anthracis* by two sgRNAs in tandem. The sterilization efficiency is no different from that of a single sgRNA. This result is similar to our previous results of using tandem sgRNA to simultaneously curing the *B. anthracis* plasmids pXO1 and pXO2. In the tandem sgRNA used for sterilization, only the second sgRNA might have an effect on sterilization, and the first sgRNA cannot form the required structure to bind Cas9, so it has no effect on killing bacteria.

Thus, our results have shown that the CRISPR-Cas9 system can be useful for gene editing in *B. cereus* group strains, and that

it can cure plasmids simply and efficiently. The cure efficiency might differ depending on the N20 target sequences that are chosen. Our results indicate that the plasmid elimination rate is only slightly higher when the replication initiation region is the sgRNA target, compared with the non-replication initiation region. CRISPR targeting of virulence genes can select for the loss of CRISPR function during infection, when the acquisition of those genes is under strong selective pressure (Jiang et al., 2013; Gomaa et al., 2014). Mutations in the replication initiation region have a more negative effect than mutations in the non-replication replication region, so we speculate that the Cas9 system has higher levels of off-target mutagenesis in the non-replication initiation region than in the replication initiation region, making the cure efficiency of the non-replication initiation region lower. With a lower efficiency of chromosome breakage, we speculate that off-target mutagenesis of chromosome and the self-repair mechanism in *B. anthracis* plays a key role in bacterial survival when faced with an external killing pressure. On the other side, it is also likely that the curing site we chose may not be suitable. The high efficiency of plasmid curing may only cause a change in pathogenicity, antibiotic resistance, or other metabolic processes when the selection pressure is relatively small. Furthermore, we only induced one generation of bacteria in our experiment, and increasing the induction period may increase the killing rate.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

DW, XL, and HW designed the research study. XW, YL, SiW, QZ, EF, LZ, and CP performed all the experiments. XW, DW, and XL analyzed the data. XW, XL, ShW, and DW wrote the manuscript. All authors have reviewed the final manuscript.

FUNDING

This work was supported by the State Major Science and Technology Special Projects of China (Grant No. 2018ZX10714002), the National Natural Science Foundation of China (Grant Nos. 81571958, 81601739, 81671979, and 81871619), and research funding from the State Key Laboratory of Pathogen and Biosecurity (Grant No. SKLPBS1822).

ACKNOWLEDGMENTS

We thank Professor Jun Cai from Nankai University for gifting the pJOE8999 plasmid. We thank Sandra Cheesman, Ph.D., from Liwen Bianji, Edanz Group China (www.liwenbianji.cn), for editing the English text of a draft of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.536357/full#supplementary-material

REFERENCES

Agata, N., Ohta, M., and Mori, M. (1996). Production of an emetic toxin, cereulide, is associated with a specific class of *Bacillus cereus*. *Curr. Microbiol.* 33, 67–69. doi: 10.1007/s002849900076

Agathe, B., Melissa, D., Hilsenbeck, S. G., Peterson, S. N., and Koehler, T. M. (2003). Global effects of virulence gene regulators in a *Bacillus anthracis* strain with both virulence plasmids. *Infect. Immun.* 71, 2736–2743. doi: 10.1128/IAI.71.5.2736-2743.2003

Altenbuchner, J. (2016). Editing of the *Bacillus subtilis* genome by the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 82:AEM.01453–01416. doi: 10.1128/AEM.01453-16

Ariel, N., Zvi, A., Makarova, K. S., Chitlaru, T., Elhanany, E., Velan, B., et al. (2003). Genome-based bioinformatic selection of chromosomal *Bacillus anthracis* putative vaccine candidates coupled with proteomic identification of surface-associated antigens. *Infect. Immunity* 71, 4563–4579. doi: 10.1128/IAI.71.8.4563-4579.2003

Arnesen, L. P. S., Fagerlund, A., and Granum, P. E. (2008). From soil to gut: *Bacillus cereus* and its food poisoning toxins. *Fems Microbiol. Rev.* 32, 579–606. doi: 10.1111/j.1574-6976.2008.00112.x

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. Biotech.* 32, 1146–1150. doi: 10.1038/nbt.3043

Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during *in vivo* bacterial infection. *Cell Host Microbe* 12, 177–186. doi: 10.1016/j.chom.2012.06.003

Citorik, R. J., Mimee, M., and Lu, T. K. (2014). Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotech.* 32, 1141–1145. doi: 10.1038/nbt.3011

Drobniewski, F. A. (1993). Bacillus cereus and related species. *Clin. Microbiol. Rev.* 6, 324–338. doi: 10.1128/CMR.6.4.324

Ehling-Schulz, M., Vukov, N., Schulz, A., Shaheen, R., Andersson, M., Martlbauer, E., et al. (2005). Identification and partial characterization of the nonribosomal peptide synthetase gene responsible for cereulide production in emetic *Bacillus cereus*. *Appl. Environ. Microbiol.* 71, 105–113. doi: 10.1128/AEM.71.1.105-113.2005

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. doi: 10.1128/mBio.00928-13

Hauge, S. (1950). *Food Poisoning Caused by Bacillus cereus*. A Preliminary Report. Nordisk Hygienisk Tidskrift.

Helgason, E. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringensis*-one species on basis of genetic evidence. *Appl. Environ. Microbiol.* 66, 2627–2630. doi: 10.1128/AEM.66.6.2627-2630.2000

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

Levy, H., Weiss, S., Altboum, Z., Schlomovitz, J., Glinert, I., Sittner, A., et al. (2012). Differential contribution of bacillus anthracis toxins to pathogenicity in two animal models. *Infect. Immunity* 80, 2623–2631. doi: 10.1128/IAI.00244-12

Liu, X., Wang, D., Wang, H., Feng, E., Li, Z., and Wang, H. (2012). Curing of plasmid pXO1 frombacillus anthracisusing plasmid incompatibility. *PLoS ONE* 7:e29875. doi: 10.1371/journal.pone.0029875

Novick, R. P., Clowes, R. C., Cohen, S. N., Curtiss, R., Datta, N., and Falkow, S. (2009). Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol Rev.* 40:168–189. doi: 10.1128/BR.40.1.168-189.1976

Novick, R. P., and Hoppensteadt, F. C. (1978). On plasmid incompatibility. *Microbiol. Rev.* 1, 421–434. doi: 10.1016/0147-619X(78)90001-X

Ramisse, V., Patra, G., Garrigue, H., Guesdon, J. L., and Mock, M. (1996). Identification and characterization of *Bacillus anthracis* by multiplex PCR analysis of sequences on plasmids pXO1 and pXO2 and chromosomal DNA. *Fems Microbiol. Lett.* 145, 9–16. doi: 10.1111/j.1574-6968.1996.tb08548.x

Shatalin, K. Y., and Neyfakh, A. A. (2005). Efficient gene inactivation in *Bacillus anthracis*. *Fems Microbiol. Lett.* 245, 315–319. doi: 10.1016/j.femsle.2005.03.029

Shinagawa, K., Konuma, H., Sekita, H., and Sugii, S. (1995). Emesis of rhesus monkeys induced by intragastric administration with the HEp-2 vacuolation factor (cereulide) produced by *Bacillus cereus*. *FEMS Microbiol. Lett.* 130, 87–90. doi: 10.1016/0378-1097(95)00188-B

Trevors, J. T. (1986). Plasmid curing in bacteria. *Fems Microbiol. Lett.* 32, 149–157. doi: 10.1111/j.1574-6968.1986.tb01189.x

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, H., Liu, X., Feng, E., Li, Z., Wang, D., Liao, X., et al. (2011). Curing the plasmid pXO2 from *Bacillus anthracis* A16 using plasmid incompatibility. *Curr. Microbiol.* 62, 703–709. doi: 10.1007/s00284-010-9767-2

Weapon, A. A. (2002). Anthrax as a biological Weapon, 2002: updated recommendations for management—correction. *J. Am. Med. Assoc.* 288:1849. doi: 10.1001/jama.287.17.2236

Wui, S. R., Ji, E. H., Kim, Y. H., Rhie, G. E., and Na, G. L. (2013). Increased long-term immunity to *Bacillus anthracis* protective antigen in mice immunized with a CIA06B-adjuvanted anthrax vaccine. *Arch. Pharmacal Res.* 36, 464–471. doi: 10.1007/s12272-013-0034-5

Xie, S., Bin, S., Chaobao, Z., Xingxu, H., Yonglian, Z., B., et al. (2014). sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS ONE* 9:e100448. doi: 10.1371/journal.pone.0100448

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership